

1 31
TCG ACT ATG AAT GCT GAT ACT GCT CCA ACA TCT CCT TGT CCT TCC ATA TCT TCC CAG AAC
ser thr met asn ala asp thr ala pro thr ser pro cys pro ser ile ser ser gln asn

61 91
TCA AGC TCC TGC TCC AGC TTC CAG GAC CAG AAG ATC GCC AGC ATG TTC GAT CGG ACT TCC
ser ser ser cys ser ser phe gln asp gln lys ile ala ser met phe asp arg thr ser

121 Cadherin
151 |xx EC motif xx|
AGA GTA CCC GCC AGC AGC ACT TCC TCA CCG GGG CTC CTC TTC ACA GAA CTG GCT GCT GCC
arg val pro ala ser ser thr ser ser pro gly leu leu phe thr glu leu ala ala ala

181 211
CTG GAT GCC GAA GGG GAA GGA ATC AGC GAA GTA CAA AGG AAA GCT GTC AGT GCA ATT CAC
leu asp ala glu gly glu gly ile ser glu val gln arg lys ala val ser ala ile his

241 271
AGC CTG CTA AGT TCT CAC GAC CTG GAC CCA CGC TGT GTC AAA CCA GAG GTG AAG GTC AAA
ser leu leu ser ser his asp leu asp pro arg cys val lys pro glu val lys val lys

301 331
ATC GCC GCC CTT TAC CTA CCT TTA GTT GGC ATC ATT TTG GAT GCT TTG CCA CAG CTC TGT
ile ala ala leu tyr leu pro leu val gly ile ile leu asp ala leu pro gln leu cys

361 391
GAC TTT ACA GTT GCA GAT ACT CGC AGA TAC CGC ACC AGT GGC TCG GAT GAA GAA CAA GAA
asp phe thr val ala asp thr arg arg tyr arg thr ser gly ser asp glu glu gln glu

421 451
GGA GCC GGT GCC ATT ACC CAG AAT GTG GCT CTG GCC ATA GCA GGG AAT AAT TTC AAT TTG
gly ala gly ala ile thr gln asn val ala leu ala ile ala gly asn asn phe asn leu

481 511
AAA ACA AGT GGA ATA GTG CTG TCT TCC TTG CCC TAT AAG CAG TAC AAC ATG CTG AAC GCG
lys thr ser gly ile val leu ser ser leu pro tyr lys gln tyr asn met leu asn ala

541 571
GAC ACT ACT CGC AAC CTC ATG ATC TGC TTC CTC TGG ATC ATG AAA AAT GCT GAT CAG AGC
asp thr thr arg asn leu met ile cys phe leu trp ile met lys asn ala asp gln ser

601 631
CTC ATT AGG AAG TGG ATT GCT GAC CTG CCA TCA ACG CAG CTC AAC AGG ATT TTA GAT CTA
leu ile arg lys trp ile ala asp leu pro ser thr gln leu asn arg ile leu asp leu

661

691

CTT TTC ATC TGT GTG TTA TGT TTT GAG TAT AAG GGA AAA CAG AGT TCT GAC AAA GTC AGT
leu phe ile cys val leu cys phe glu tyr lys gly lys gln ser ser asp lys val ser

721

751

ACC CAA GTC CTG CAG AAG TCA AGG GAT GTC AAG GCC CGG CTG GAA GAG GCT TTG CTG CGT
thr gln val leu gln lys ser arg asp val lys ala arg leu glu glu ala leu leu arg

781

811

GGG GAA GGG GCC AGA GGG GAG ATG ATG CGC CGC CGG GCT CCA GGG AAC GAC CGA TTT CCA
gly glu gly ala arg gly glu met met arg arg arg ala pro gly asn asp arg phe pro

841

871

GGC CTA AAT GAA AAT TTG AGA TGG AAG AAA GAG CAG ACA CAT TGG CGG CAA GCT AAT GAG
gly leu asn glu asn leu arg trp lys glu gln thr his trp arg gln ala asn glu

901

931

AAG CTA GAT AAA ACA AAG GCC GAG TTA GAT CAA GAA GCC TTG ATC AGT GGC AAT CTG GCT
lys leu asp lys thr lys ala glu leu asp gln glu ala leu ile ser gly asn leu ala

961

991

ACA GAA GCA CAT TTA ATC ATC CTG GAT ATG CAG GAA AAC ATT ATC CAG GCG AGC TCG GCT
thr glu ala his leu ile ile leu asp met gln glu asn ile ile gln ala ser ser ala

1021

1051

CTG GAC TGT AAA GAC AGC CTG CTG GGA GGT TTG CTG AGG GTG CTG GTG AAT TCT CTG AAC
leu asp cys lys asp ser leu leu gly gly val leu arg val leu val asn ser leu asn

1081

1111

TGT GAT CAG AGT ACC ACC TAC CTG ACT CAC TGC TTT GCA ACA CTC CGT GCT CTC ATC GCC
cys asp gln ser thr thr tyr leu thr his cys phe ala thr leu arg ala leu ile ala

1141

1171

AAG TTT GGA GAC TTA CTC TTC GAA GAG GAG GTG GAA CAG TGT TTC GAC CTA TGT CAC CAA
lys phe gly asp leu leu phe glu glu val glu gln cys phe asp leu cys his gln

1201

1231

GTC CTG CAC CAC TGC AGC AGC AGC ATG GAT GTC ACC CGG AGC CAA GCC TGT GCC ACC CTT
val leu his his cys ser ser ser met asp val thr arg ser gln ala cys ala thr leu

1261

1291

TAC CTC CTC ATG AGG TTC AGT TTT GGA GCC ACC AGT AAT TTT GCA AGA GTA AAG ATG CAA
tyr leu leu met arg phe ser phe gly ala thr ser asn phe ala arg val lys met gln

1321

1351

GTA ACC ATG TCC CTG GCA TCT TTG GTG GGA AGA GCA CCA GAC TTT AAT GAA GAG CAC CTG
val thr met ser leu ala ser leu val gly arg ala pro asp phe asn glu glu his leu

1381

AGA AGA TCC TTG AGG ACA ATT TTG GCC TAT TCA GAA GAG GAC ACA GCC ATG CAG ATG ACT
arg arg ser leu arg thr ile leu ala tyr ser glu glu asp thr ala met gln met thr

1441

CCT TTT CCC ACC CAG GTG GAG GAA CTT CTC TGT AAT CTG AAT AGC ATC TTA TAT GAC ACA
pro phe pro thr gln val glu glu leu leu cys asn leu asn ser ile leu tyr asp thr

1501

GTG AAA ATG AGG GAA TTT CAG GAA GAT CCT GAG ATG CTT ATG GAT CTC ATG TAC AGA ATT
val lys met arg glu phe gln glu asp pro glu met leu met asp leu met tyr arg ile

1561

GCC AAG AGT TAC CAG GCA TCT CCT GAT CTG CGG CTG ACC TGG CTC CAG AAC ATG GCA GAG
ala lys ser tyr gln ala ser pro asp leu arg leu thr trp leu gln asn met ala glu

1621

|xxxxxxxxxxxxxxxxxxxxxx transmembrane domain xxxxxxxx
AAA CAC ACC AAG AAG AAG TGC TAC ACG GAG GCT GCC ATG TGC GTG CAC GCC GCT GCG
lys his thr lys lys cys tyr thr glu ala ala met cys leu val his ala ala ala

xxxxxxxxxxxxxxxxxxxxxx|

1711

TTA GTG GCT GAG TAT CTG AGC ATG CTG GAG GAC CAC AGC TAC CTG CCC GTG GGC AGT GTC
leu val ala glu tyr leu ser met leu glu asp his ser tyr leu pro val gly ser val

1741

AGC TTC CAG AAT ATT TCT TCC AAT GTG CTG GAG GAG TCT GTG GTC TCT GAG GAC ACC CTG
ser phe gln asn ile ser ser asn val leu glu glu ser val val ser glu asp thr leu

1801

TCA CCT GAC GAG GAT GGG GTG TGC GCA GGC CAG TAC TTC ACC GAG AGT GGC CTG GTA GGC
ser pro asp glu asp gly val cys ala gly gln tyr phe thr glu ser gly leu val gly

1861

CTC CTG GAG CAG GCC GCG GAG CTC TTC AGC ACG GGA GGC TTA TAT GAG ACA GTT AAT GAG
leu leu glu gln ala ala glu leu phe ser thr gly gly leu tyr glu thr val asn glu

1921

GTC TAC AAG CTG GTC ATC CCC ATC CTA GAA GCG CAT CGA GAA TTC CGG AAG CTG ACA CTC
val tyr lys leu val ile pro ile leu glu ala his arg glu phe arg lys leu thr leu

1981

ACT CAC AGC AAG CTG CAG AGA GCC TTC GAC AGC ATC GTT AAC AAG GAT CAT AAG AGA ATG
thr his ser lys leu gln arg ala phe asp ser ile val asn lys asp his lys arg met

2041

|xxxxx ITAM xxxx| 2071
TTT GGA ACC TAC TTC CGA GTT GGT TTC TTT GGA TCC AAA TTT GGG GAT TTG GAT GAA CAG
phe gly thr tyr phe arg val gly phe phe gly ser lys phe gly asp leu asp glu gln

2101 2131
GAG TTT GTC TAC AAA GAG CCT GCA ATT ACC AAG CTT CCT GAG ATC TCA CAT AGA CTA GAG
glu phe val tyr lys glu pro ala ile thr lys leu pro glu ile ser his arg leu glu

2161 2191
GCA TTT TAT GGT CAA TGT TTT GGT GCA GAA TTT GTG GAA GTG ATT AAA GAC TCC ACT CCT
ala phe tyr gly gln cys phe gly ala glu phe val glu val ile lys asp ser thr pro

2221 2251
GTG GAC AAA ACC AAG TTG GAT CCT AAC AAG GCC TAC ATA CAG ATC ACT TTT GTG GAG CCC
val asp lys thr lys leu asp pro asn lys ala tyr ile gln ile thr phe val glu pro

2281 2311
TAC TTT GAT GAG TAT GAG ATG AAA GAC AGG GTC ACA TAC TTT GAG AAG AAT TTC AAC CTC
tyr phe asp glu tyr glu met lys asp arg val thr tyr phe glu lys asn phe asn leu

2341 2371
CGG AGG TTC ATG TAC ACC ACC CCG TTC ACC CTG GAG GGG CGG CCT CGG GGA GAG CTG CAT
arg arg phe met tyr thr pro phe thr leu glu gly arg pro arg gly glu leu his

2401 2431
GAG CAG TAC AGA AGG AAC ACA GTC CTG ACC ACT ATG CAC GCC TTC CCC TAC ATC AAG ACC
glu gln tyr arg arg asn thr val leu thr thr met his ala phe pro tyr ile lys thr

2461 2491 |xxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
AGG ATC AGC GTC ATC CAG AAG GAG GAG TTT GTT TTG ACA CCG ATT GAA GTT GCC ATT GAA
arg ile ser val ile gln lys glu glu phe val leu thr pro ile glu val ala ile glu
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx Coiled coil 1 xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
GAC ATG AAG AAG ACC CTG CAG TTA GCA GTT GCC ATT AAC CAG GAG CCG CCT GAT GCA
asp met lys lys thr leu gln leu ala val ala ile asn gln glu pro pro asp ala

xxxxxxxxxxxxxxxxxxxxxxxxxxxxxx| 2611
AAG ATG CTT CAG ATG GTG CTG CAA GGC TCT GTG GGA GCT ACT GTA AAT CAG GGA CCA CTG
lys met leu gln met val leu gln gly ser val gly ala thr val asn gln gly pro leu

2641 2671
GAA GTA GCC CAA GTG TTT TTG GCT GAA ATT CCT GCT GAT CCA AAA CTC TAT CGA CAT CAC
glu val ala gln val phe leu ala glu ile pro ala asp pro lys leu tyr arg his his

2701 2731 |xxxxxxxxxxxxx
AAC AAG TTG AGG TTA TGC TTT AAG GAA TTC ATC ATG AGA TGT GGT GAA GCT GTA GAG AAA
asn lys leu arg leu cys phe lys glu phe ile met arg cys gly glu ala val glu lys

xxxxxxxxxxxxxxxxxxxxxxxxxxxxxx Coiled coil 2 xxxxxxxxxxxxxxxxxxxxxxxxx
AAC AAG CGT CTC ATC ACG GCA GAC CAG AGG GAA TAT CAG CAG GAA CTC AAA AAG AAC TAT
asn lys arg leu ile thr ala asp gln arg glu tyr gln gln glu leu lys asn tyr

xxx|
AAC AAG CTA AAA GAG AAC CTC AGG CCA ATG ATC GAG CGG AAA ATT CCA GAA CTG TAC AAG
asn lys leu lys glu asn leu arg pro met ile glu arg lys ile pro glu leu tyr lys

2881 2911
CCA ATA TTC AGA GTT GAG AGT CAA AAG AGG GAC TCC TTC CAC AGA TCT AGT TTC AGG AAA
pro ile phe arg val glu ser gln lys arg asp ser phe his arg ser ser phe arg lys

2941 2971
TGT GAA ACC CAG TTG TCA CAG GGC AGC TAA GAA AAG CCA TCT TCA TTC GTG GAG ACT GTG
cys glu thr gln leu ser gln gly ser OCH glu lys pro ser ser phe val glu thr val

3001 3031
GCC CTG CAA CCC TGG AGA AGG ACT TGC TGG TAC TTA AAA AAT GGG ACA TTT GCC ACC CAG
ala leu gln pro trp arg arg thr cys trp tyr leu lys asn gly thr phe ala thr gln

3061 3091
GAC TGA CTG TAC ACT CCC TGA TCA GCC AGC ACT CTG GAA GCT TTG GGA TCC CAG GAA CCA
asp STP

3121 3151
TGG AAT TAT TCC CAA ATG GAC TCT GAC CAG ATT TTT GCC ATA CTG GGG GGT GGC GGG ATG

3181 3211
GAG GAT GGG TAC TCA GGC ATG ACT GCG TAT TTA TTA AAG TGT GTT TTT CCA CAA TGT ACC

3241 3271
AAA CAA GGC ATA AGC AGC TTC TCC TGC TGA CTG GCC AAT CAC TGC CCA TCT GAG AGA TGA

3301 3331
TTT CCT CTG GCC CAT ATT TGA ATT TAT TGG AGT AAC TCA AAT TGC CTG AGG AAA AAT GGA

3361 3391
AAA ATT ATC CAC CAG TCG ATT CAA ACT GAA TTT CAC TCT TTA TAG GAA GGC AGG GCA AAC

3421 3451
TTG TAG GAG TAC GAA ACA TTT TCA ATA AAT CTA CAA AGG GAA GCC TTA CTA CAA TTC CAA

3481 3511
AAA TCA TCA TGG TTG GAA ATT TGG GAG GAG ATT ATT TGT GAA CTT GTT ACC CTT TTG GTA

3541 3571
ATG GTG GAC TAA TTG CTG TAT AGT TAT TTT TGT TTT ATT ATT ACT GTT ACA TTA ATT TAA

3601 3631
CAT GCA TTT ATA GAA GAA TAC ATT CAA AGC ACT GAT GTA GGA GAT ACA CGG TAC TTG GAG

3661 3691
.....

3721 3751
CTT TGC TTT TTT TCT TAT GTC ACT CTT GTG TAC TAT CTA TTT TTC TCC TCT CTG GGA CCA

3781 3811
AGT TTC TTT TTA TAA AGC AAT AAT ATC TCT GTT TTC ATT TCA GAA CAT TGT GCT GTC TGT

3841 3871
CAG CAT ATG TAT ATC AGC TAC AAA ATA TAT TCA ACT TTG ACT TCT TTT GAC AAA GGA CTT

3901 3931
TAG GAA AAG GAG GAA CAA AGA CAT TAT TTG AGA ATT AAA TTA TAT ATT TTT AAT ATG ACT

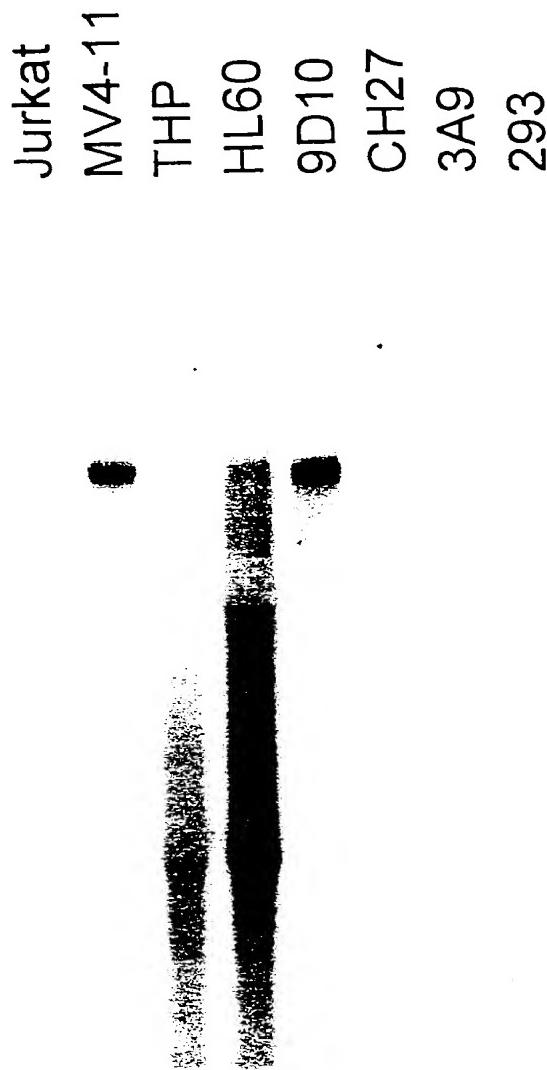
3961 3991
GTG ACC TTG ACT GAT AAT AAA GAT GTA ATA AGA ATT GCA AGC TAA AAA AAA AAA AAA AAA

4021 AAC TCG

BRAIN
HEART
SKEL. MUSCLE
COLON
THYMUS
SPLEEN
KIDNEY
LIVER
SM. INTESTIN
PLACENTA
LUNG
PBL

.5 kb —

A



B

Fig. 2

HC2A
KIAA
rat
HC4
HC1
HC3
HC5

ASGNLDKNARFSAIYRQDSNKLSDNDMLKLLADFRKPEKMAKLPVILGNLDITIDNVSD

HC2A
KIAA
rat
HC4
HC1
HC3
HC5

FPNYVNSSYIPTKQFETCSKTPITFEVEEFVPCIPKHTQPYTIYTNHLYVYPKYLKYDSQ

HC2A
KIAA
rat
HC4
HC1
HC3
HC5

KSFAKARNIAICIEFKDSDEEDSQPLKCIYGRPGGPVFTRSAAVLHHQNPEFYDEIK

HC2A
KIAA
rat
HC4
HC1
HC3
HC5

IELPTQLHEKHHLLLTFHVSCDNSSKGSTKKRDVVTQVGYSWLPLLKDGRVVTEQH
IELPTQLHEKHHLLLTFHVSCDNSSKGSTKKRDVVTQVGYSWLPLLKDGRVVTEQH

HC2A
KIAA
rat
HC4
HC1
HC3
HC5

PVSANLPSGYLGQELGMGRHYGPEIKWVDGGKPLLKISTHLVSTVYTQDQHLHNFFQYC
PVSANLPSGYLGQELGMGRHYGPEIKWVDGGKPLLKISTHLVSTVYTQDQHLHNFFQYC

HC2A
KIAA
rat
HC4
HC1
HC3
HC5

GPGPARSTVSISLISNSARV

HC2A
KIAA
rat
HC4
HC1
HC3
HC5

QKTESGAQALGNELVKYLKSLHAMEGHVMIAFLPTILNQLFRVLT-RATQEEVAVNVTRV
QKTESGAQALGNELVKYLKSLHAMEGHVMIAFLPTILNQLFRVLT-RATQEEVAVNVTRV
-----MEIQVLIIRFLSVIIMQLFWVLPNMIHEDDVPISCP
-----MSFLPIIILNQLFKVLV-QNEEDEITTTVTRV
NRSRSLSNSNPDISGPTSPDDEVRSIIIGSKGLDRSNWNTGGPKAAPWGSNPSPSAES

HC2A	I I HVVAQC [REDACTED] LESHLR SYV KYAY KAE PYV ASE YKTVHEELTH [REDACTED] TILKPSADFLTSN
KIAA	I I HVVAQC [REDACTED] LESHLR SYV KYAY KAE PYV ASE YKTVHEELTH [REDACTED] TILKPSADFLTSN
rat	[REDACTED]
HC4	LFHIVSKCHEEGLDSYLSFIKYSFRPGKPSAPQAPIHETLATMMIALLKQSADFLAIN
HC1	L P D I V A K C H E E Q L D H S V Q S Y I K F V F K T R -- A C K E R P V H E D L A K N V T G L K - S N D S P T V K
HC3	T Q A M D R S C N R M S S H T E T S S F L Q T L T G R L P -- T K K L F H E E L A L Q W V V C S G -- S V R -- E
HC5	[REDACTED]
Cadherin Cleavage	
HC2A	KLLRYSWFFF DV LIKSMAQHLIENS KVKLI [REDACTED] RNQRF PASYHAAE TVVNMLMPHITQKFGD
KIAA	KLLKYSWFFF DV LIKSMAQHLIENS KVKLI [REDACTED] RNQRF PASYHAAE TVVNMLMPHITQKFRD
rat	[REDACTED]
HC4	KLLKYSWFFF EI IA KSMA TYLLEEN KIKLTHGQRFPKAYHHALHSLFLAIT- I V E S Q Y A E
HC1	HVLKHSWFFF AI ILKSMAQHLIDTNKIQLEPRPQRFPESYQN ELDNLVMVLSDHVIWKYKD
HC3	SALQQAWFFF ELMVKS MVHHLYFNDKLEPRKSRFPER FMDDIAALVSTIASDIVSRFQK
HC5	[REDACTED]
HC2A	N P E A S K N A N H S L A V F I K R C F T F M D R G F V F K Q I N -- N Y I S -- C F A P G D P K T L F E Y K F E F L
KIAA	N P E A S K N A N H S L A V F I K R C F T F M D R G F V F K Q I N -- N Y I S -- C F A P G D P K T L F E Y K F E F L
rat	[REDACTED]
HC4	I P K E S R N V N Y S I A S F L K C C L T L M D R G F V F N L I N -- D Y I S -- G F S P K D P K V L A E Y K F E F L
HC1	A L E E T R R A T H S V A R F L K R C F T F M D R G C V F K M V N -- N Y I S -- M F S S G D L K T L C Q Y K F D F L
HC3	D T E M V E R L N T S L A F F L N D L L S V M D R G F V F S L I K S C Y K Q V S S K L Y S L P N P S V L V S L R L D F L
HC5	[REDACTED]
HC2A	F V V C N H E H Y I P L N L P M -- P F G K G R I Q R -- Y Q D L Q L -- D Y S L T D E F
KIAA	F V V C N H E H Y I P L N L P M -- P F G K G R I Q R -- Y Q D L Q L -- D Y S L T D E F
rat	[REDACTED]
HC4	C T I C N H E H Y I P L N L P M -- A F A K P K L Q R -- V Q D S N L -- E Y S L S D E Y
HC1	Q E V C Q H E H F I P L C L P I R S A N I P D P L T P S E S -- T Q E L H A S D M P E Y S V T N E F
HC3	F I I C S H E H Y V T L N L P C S L L T P P A S P S P S V S S A T S Q S S G F S T N V Q D Q K I A N M F E L S -- V P F
HC5	M N A L T A P T S P C P S I S -- S Q N S S C S S F Q D Q K I A S M F D R T S R V P A
Cadherin EC motif	
HC2A	CRNHFLVGILLREVGTLQE FRE --- VRLIAISVLKNLLIKHSFDDRYASRSHQARIAT
KIAA	CRNHFLVGILLREVGTLQE FRE --- VRLIAISVLKNLLIKHSFDDRYASRSHQARIAT
rat	[REDACTED]
HC4	C K H H F L V G I L L R E T S I A L Q D N Y E -- I R Y T A I S V I K N L L I K H A F D T R Y Q H K N Q Q A K I A Q
HC1	C R K H F L I G I L L R E V G F A L Q E D Q D -- V R H L A L A V L K N L M A K H S F D D R Y R E P R K Q A Q I A S
HC3	F QQHYLAGI VLT E L A V I L L P D A E G L F G L H K K V I N M V H N L L S H D S D P R Y S D P Q I K A R V A M
HC5	S S T S - S P G I L F T E L A A A L D A E G E G I S E V Q R K A V S A I H S L L S H D L D P R C V K P E V K V K I A A
HC2A	LYLPLFGLL I ENVQR I N V F D V S P F P V N A G - M T V K D E S L A L P A V N P L V T P Q K G S T L D N S L H
KIAA	LYLPLFGLL I ENVQR I N V F D V S P F P V N A G - M T V K D E S L A L P A V N P L V T P Q K G S T L D N S L H
rat	[REDACTED]
HC4	LYL P F V G L L E N I Q R L A G F D T L Y S C A A M P N S A R D E F P C G -- F T S P -- A N -- R G S L S
HC1	LYM P L Y G M L L D N M P R I Y L K D L Y P F T V N T S N Q G S R D D L S T N G G F Q S Q T A I K H A N S V D T S F S
HC3	LYL P L I G I I M E T V P Q L Y D F T E T H N Q R G R P C I A T D D Y E S E -- S G -- S M I S
HC5	LYL P L V G I I L D A L P Q L C D F T V A D T R R Y R -- T S G S D E E Q E -- G A -- G A I T

A

HC2A KDLLGAISC [SPYTTSTPNINSVRNADSRGSLISTDSGNSLPE] SEKNSLDKHQQSS
KIAA KDLLGAIS [SPYTTSTPNINSVRNADSRGSLISTDSGNSLPE] SEKNSLDKHQQSS
rat -----
HC4 TDKD TAYGSFQNG-----HG I KREDSRGSLIP-EGATGFPDQGNTGEN-----TRQS
HC1 KDVLNSIAFSS-----IAI STVNHADSRASLASLDSNPSTNEKSSEKTNDCEKIPRPL
HC3 QTVAMAIAGTSPQ-----LTRPGSFLLTSTSQRQHT-----
HC5 QNVALAIAGNNFN-----LKTS G-IVLSSL PYKQYN-----

HC2A TLGNSVVRCDKLDQSEIKSLLMCFLYILKSMDDALFTYWN-KASTSELMDFFTISEVCL
KIAA TLGNSVVRCDKLDQSEIKSLLMCFLYILKSMDDALFTYWN-KASTSELMDFFTISEVCL
rat -----
HC4 STRSSVSQYNRLDQYEIRSLLMCYLIVK MISEDTLLTYWN-KVSPQELINILILLEVCL
HC1 ALIGSTLRFDRLDQAETRSLLMCFLHIMKTISYETLIAYWQ-RAPSPEVSDFFSILDVCL
HC3 -----TFSAESRSLLICLWLWLN-ADETV LQKWF DLSVLQLNRLLLYLCV
HC5 -----MLNADTRNLMICFLWIMRN-ADQSLIRKWIADLPSTQLNRILDL FICV

HC2A HQFQYMGKRYIARNQEGLG--PIVHDRKS-----QTLPVSRNRTGMM
KIAA HQFQYMGKRYIAR-----TGMM
rat -----
HC4 FHF RYMGKRNIA RVHD A WL SKHFG IDR KS-----Q TMPALRN RSGVM
HC1 QNF RYLGKRNIA R KIAA AF--KF VQ STQ NNG TL KG S NP SC QT SG LLA QWM H ST SR HE GH K
HC3 SC F EYKGK KV FER MN SLTFK--KSKDMRAK-----LEEAI LG SIGAR QEM V
HC5 LCF EYKGK QSS DKV STQ VL Q--KS RDVK AR-----LEEAL LR GEGAR GEM M

HC2A HARLQQL-----GSL DNS-----LTF NH SYGH SD AD VLHQ S LLE ANI ATE VC
KIAA HARLQQL-----GSL DNS-----LTF NH SYGH SD AD VLHQ S LLE ANI ATE VC
rat -----
HC4 QARLQHL-----SSLESS-----FTL NH S STT EADI FH Q ALLEG NTATE VS
HC1 QH RS Q TLP I IRG K---N A LS NP K L---LQ M LD N TMT S NS N EID IV HH VD TEAN I ATE EG C
HC3 RRS RG Q LER SP SG SA FG S QEN LR W R K DM TH W R Q N TE K L D K S R A E I E H E A L I D G N I ATE A N
HC5 RRRAPGND R FP-----G LN EN LR W KKE Q TH W R Q A N E K L D K T K A E L D Q E A L I S G N I ATE A H

HC2A LTALDTLSLFTLAFKNQLLADHGHNPLMKKVFDVYLCFLQKHQSETALKNVFTALRS LI Y
KIAA LTALDTLSLFTLAFKNQLLADHGHNPLMKKVFDVYLCFLQKHQSETALKNVFTALRS LI Y
rat -----
HC4 LTV LDT IS FFT QCF KTHFLNN DGHNPLM KK VFDI HLAFLKNGQSEVSLKH VFA S LRA FIS
HC1 LTI LDDLVLSLFTQTHQRQLQQCDCQNSLMKRGFD TYMLFFQVNQS AT AL KH VFA S LRL FVC
HC3 LII LD TLE I VV QTV S--V TES--KESI LGGV LKV LLHS MACN QSAV YLQ HC F AT QRAL VS
HC5 LII LD M QEN II QASS--ALDC--KDS L LGGV LRV LV NSL NC QST T YLTHC F AT L RAL IA

HC2A KFPSTFYEGRADMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKS FVRTH
KIAA KFPSTFYEGRADMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKS FVRTH
rat KFPSTFYEGRADMCAALCYEILKCCNSKLSSIRTEASQLLYFLMRNNFDYTGKKS FVRTH
HC4 KFP S A FF KGR VNM CAA FCYEV LKCC TSK I S S TRNE AS ALLY LLM RNN FEY T KR KT FLR TH
HC1 KFP S A FF QGP ADL CGS F CYEV LKCC NHR SR ST QTE AS ALLY L FMR KN FEF NK QKS I VR SH
HC3 KFP ELL FEE ETE QCA DL CL RLL RH C S S I G T I RSH P S A S LY LLM RQ NF E I G N --N FAR VK
HC5 KFG D L L FEE E VE QCF DL CH QV L HHC S S S MD VTR S QAC AT LY LLM RFS FG AT S --N FAR VK

HC2A LQVIISVSQLIADVVGIGGETRFQQLSIIINNCANS DR LIKHTSFSSDV KDL T K R I RTV LM
KIAA LQVIISVSQLIADVVGIGGTRFQQLSIIINNCANS DR LIKHTSFSSDV KDL T K R I RTV LM
rat LQVIISLSQLIADVVGIGGTRFQQLSIIINNCANS DR LIKHTSFSSDV KDL T K R I RTV LM
HC4 LQIII AVSQLIADVVALSGGSRFQESLFIIINNFANS DR PM LARAFPAEV KDL T K R I RTV LM
HC5 LQIII AVSQLIADVVALSGGSRFQESLFIIINNFANS DR PM LARAFPAEV KDL T K R I RTV LM

Transmembrane

HC2A	ATAQMKEHENDPEMLVLDLQYSI	LAKS YAST PELRKTWLDSMARI	HVKNGDLSEAAMCYVHV
KIAA	ATAQMKEHENDPEMLVLDLQYSI	LAKS YAST PELRKTWLDSMARI	HVKNGDLSEAAMCYVHV
rat	ATAQMKEHENDPEMLVLDLQYSI	LAKS YAST PELRKTWLDSMARI	HVKNGDLSEAAMCYVHV
HC4	ATAQMKEHEKDPEMLIDLQYSI	LAKS YAST PELRKTWLDSMAKI	HVKNGDFSEAAMCYVHV
HC1	ATAQMKEHEKDPEMLIDLQYSI	LANS YAST PELRRTWLESMAKI	HARNGDLSEAAMCYIHI
HC3	DTVKMKEHQEDPEMLIDILMYRI	AKGYQTSPDLRLTWLQNMA	GKHSERSNHAEEAQCLVHS
HC5	DTVKMREFQEDPEMLIDILMYRI	AKSYQASPDRLRTWLQNMA	EHTKKCYTEAMCLVHS

	domain	SH3		
HC2A	TALVAEYI	TRKGV-----	-FRQGCTAFRVITPN	
KIAA	TALVAEYI	TRKEA-----	-VQWEPPPLLPHSHSACLRRSRGGVFRQGCTAFRVITPN	
rat	TALVAEYI	TRKEA-----	-LAIQREPPVFVPSHTSCQRKSRGGMFROQGCTAFRVITPN	
HC4	AALVAEFL	HRKKL-----	-FPNGCSAFKKITPN	
HC1	AALIAEYI	KRKGYWKVEKICTASLLSEDTHP	CDSNSLLTPSGGSMFSMGWP	PAFLSITPN
HC3	AALVAEYI	SMLED-----	-RKYLPVGCVTFQNISSN	
HC5	AALVAEYI	SMLED-----	-HSYLPVGGSVSFQNISSN	

	ITAM		
HC2A	I DEEASMMEDVGMD-----	VHFNEDVLMELLEQCADGLWKAER	YEI LIADIYKL I IPI
KIAA	I DEEASMMEDVGMD-----	VHFNEDVLMELLEQCADGLWKAER	YEI LIADIYKL I IPI
rat	I DEEASMMEDVGMD-----	VHFNEDVLMELLEQCADGLWKAER	LRAGLTSINSSSP
HC4	I DEEGAMKEDAGMMD-----	VHYSEEVLLELCQCVNGLWKAER	YEI SEISKLIGPI
HC1	I KEEGAAKEDSGMHD-----	TPYNENILVEQLYMCGEFLWKS	ER YEL I ADVNKPIIAV
HC3	VLEESAVSDDVVSPDEEGICSGKY	FTESGLVGLLEQAAASF	SMAGMYEA
HC5	VLEESVVSEDTLPD	EDGVCAGQYFTESGLVGLLEQAAEL	FSTGGLYETVNE Y KL I IPI

	ITAM	ITAM	ITAM	ITAM
HC2A	YEKRRD	-----	-----	-----
KIAA	YEKR	DPEPLAHLI	YDTHRAYSKV	TEVMHSGRLLGTY
rat	YEKR	DPEPLAHLI	YDTHRAYSKV	TEV
HC4	YEKR	DPEPLAHLI	YDTHRAYSKV	TEV
HC1	FEKQRDFKPLSD	YLYYD	IYDTHRAYSKV	TEV
HC3	HEANF	DAKF	LSTIHGKLQEAFSKIV	HQSTGWERMFG
HC5	LEAH	PFR	TLTHSKLQRAFDS	IVNKDH
				-KRMFG-----

	ITAM	ITAM
HC2A	-FF	FEDEDGE
KIAA	YI	YI
rat	YI	YI
HC4	YI	YI
HC1	SFF	EEEDGE
HC3	YI	YI
HC5	TKFGDLDEQ	YI

	ITAM
HC2A	YI
KIAA	YI
rat	YI
HC4	YI
HC1	YI
HC3	YI
HC5	YI

	ITAM
HC2A	YI
KIAA	YI
rat	YI
HC4	YI
HC1	YI
HC3	YI
HC5	YI

Coiled-Coil 1

HC2A	IHCFPYVKKRIPVMYQHHTDLNPIEVAIDEMSKKVAELRQLCSSAEVDMIKLQLKLQ	SV
KIAA	IHCFPYVKKRIPVMYQHHTDLNPIEVAIDEMSKKVAELRQLCSSAEVDMIKLQLKLQ	SV
rat	IHCFPYVKKRIPVMYQHHTDLNPIEVAIDEMSKKVAELHQLCSSAEVDMIKLQLKLQ	SV
HC4	SNSFPYVKKRIPINCEQQINLKPIDGATDEIKDKTAELQKLCSSSTDVD	WV
HC1	SHLFPYVKRKIQVISQSSTELNPIEVAIDEMSRKVSELNOLCTMEEVDMISLQLKLQ	SV
HC3	SHAFPYIKTPVNVTKEEILTRPIEVAIEDMQKKTQELAFATHQDPADPKMLQMVLQ	SV
HC5	MHAFPYIKTRISVIQKEEFVLTRPIEVAIEDMKKTLQLAVAINQEPPDAKMLQMVLQ	SV

Coiled-Coil 2

HC2A	SVQVNAGPLAYARAFLDDTNTKRYPDNKVLLKEVFRQFVEACGQALAVNERLIKEDQLE	
KIAA	SVQVNAGPLAYARAFLDDTNTKRYPDNKVLLKEVFRQFVEACGQALAVNERLIKEDQLE	
rat	SVQVNAGPLAYARAFLDDTNTKRYPDNKVLLKEVFRQFVEACGQALAVNERLIKEDQLE	
HC4	SVQVNAGPLAYARAFLNDSQASKYPPKKVSELKDMFRKFQACSIALELNERLIKEDQVE	
HC1	SVKVNAGPMAYARAFLLEETNAKKYPDNVKLLKEIFRQFADACGQALDVNERLIKEDQLE	
HC3	GTTVNQGPLEVAQVFLSEIPSDPKLFRHHNKRLCFKDFTKRCEDALRKNSLICPVQKE	
HC5	GATVNQGPLEVAQVFLAEIPADPKLYRHHNKRLCFKEFIMRCGEAVEKNKRLITADKRE	

Coiled-Coil 2

HC2A	YQEEMKANYFEMAKELSEIMHEQICPLEEKTS-VLPNSLHI FN A ISGTPTSTMVHGM TSS	
KIAA	YQEEMKANYFEMAKELSEIMHEQLG-----	
rat	YQEEMKANYFIRKELSDIIVPRICPGEDKRATKFP AHLQRHQ DDTNKHSGSRV DQF ILS	
HC4	YHEGLKSNFRDMVKELSDIIHEQILQEDTMHSPWMSNTLHVFCAI SGTS SDRGYGS P RYA	
HC1	YQEELRSHYKDM LSELSTVMNEQITGRDDLSK---RGVDQTCTR VISKATPALPTV SISS	
HC3	YQREL LG---KLSS---PZ-----	
HC5	YQQELKKNNYKLKENLRPMIERKIP ELYKPIFRVESQKRDS FHRSS FRK CETO LS QGS Z-	

PBM

HC2A	[SSVV]-----	
KIAA	-----	
rat	CVTLPHPHVGTCFVMCKLRTTFRANHWFCQAQEEAMGNGREKEPWTVI FNSRFYRSWGK	
HC4	EVZ-----	
HC1	SAEVZ-----	
HC3	-----	
HC5	-----	

HC2A	-----
KIAA	-----
rat	VHIFF
HC4	-----
HC1	-----
HC3	-----
HC5	-----

	A	B	C
CLASP-1	YRVAFYGO:::::::GFFEEEGKEYIYKEP		
KIAA1058	FRVAFFGQAACQYQFTDSETDVEGFFEDEDGKEYIYKEP	FEDEDGKEYIYKEP	
CLASP-2		FEDEDGKEYIYKEP	
CLASP-6	FRVAFFGQ:::::::GFFEDEDGKEYIYKEP		
CLASP-4	FRVAFYGD:::::::SFFEEEDGKEYIYKEP		
DOCK180	FAVGYYGD:::::::GFPTFLRGKVFIFYRGKEYERRED		
DOCK2	FAVGYYGD:::::::GFPSTFLRNKVFIFYRGKEYERRED		
DOCK3	FRVGFYGR:::::::KFPFFLRNKEYVCRGH		
KIAA0716	FRVGFYGR:::::::KFPFFLRNKEFVCRGH		
CLASP-3	FRVGFYGT:::::::KFGDLDEQEKFVYKEP	KEY K	
CONSENSUS	F V FYG YF	Q F R	
TRG	PKLTPLSEISQRLLKLYSDKGSENVKMIQDSGKVNPKDLDKFAYIQVTHVTFFFDEKE		
CLASP-1	PKLTGLSEISQRLLKLYADKFGADNVKIIQDSNKVNPKDLDPKYAYIQVTVTFFFEEKE		
CLASP-2	PKLTPLSEISQRLLKLYSDKGSENVKMTQDSGKVNPKDLDSKYAYIQVTHVPIFFFDEKE		
CLASP-4	PKLTGLSEISLRLVKLYGEKFGTENVKKIIQDSDKVNAKELDPKYAHIQVTVVKPYFDDKE		
CLASP-3	PAITKLAEISHRLEGFYGERFGEDVVVEVIKDSNPVDCKLDPNKAQIQITYVEPYFDYE		
KIAA0716	HDIYERLEAFQQRMLNEFPHAI-----MQHANQPDETIFQAEAQYLQIYAVTPIPESQSE		
DOCK3	HDIYERLEAFQQRMLSEFPQAVA-----MQHPNHPDDAILQCDAQYLQIYAVTPIPDVDF		
DOCK2	FQMQLMTQFFNAEK-----MNTTSAPGDDVKNAPGQYIQCFTVQPVLDDEHP		
DOCK180	EYERREDFQMQLMTQFPNAEK-----MNTTSAPGDDVKNAPGQYIQCFTVQPVLDDEHP		
CONSENSUS	L L Y M F	Y Q+ V P D L E	
D		E	
CLASP-1	RTILTTSHLPYVKKRIVQVISQSSTELNPIEVAIDEMSRKVSELN		
TRG	RTILTAIHCFPYVKKRIPVMYQHHDTLNPIEVAIDEMSKVaelH		
KIAA1058	RTILTAIHCFPYVKKRIPVMYQHHDTLNPIEVAIDEMSKVaelR		
CLASP-2	RTILTAIHCFPYVKKRIPVMYQHHDTLNPIEVAIDEMSKVaelR		
CLASP-6	RTILTTAIHCFPYVKKRIPFMYQHHDTLNPIEV:HDEMSSKVAELR		
CLASP-4	RTILTTSNSPFPYVKKRIPINCEQQINLKPIDVATDEIKDKTAELQ		
CLASP-3	RTILTTSHAFPYIKTRVNVTKEEIIILTPIEVAIEDMOKKTQELA		
CLASP-5	NTVLTMMHAFPYIKTRISVIQKEEFVLTPIEVAIEDMOKKTQELA		
KIAA0716	FTSIYLVQSLPGISRWFEVEKREVVEMSPLENAIEVLENKNOQLK		
DOCK2	FTSFVVTAYKLPGILRWFEVVMHSQTTISPLENAIETMSTANEKIL		
DOCK3	FTTILTHTHSLPGISRWFEVERRELVEVSPLENAIQVVENKNQELR		
DOCK180	FTSFVVTAYKLPGILRWFEVVMHSQTTISPLENAIETMSTANEKIL		
CONSENSUS	RT L F P V + V + P+E AI+ M + L F L L	+ I	

	F	G	
CLASP-1	SEQLKLQGSVSVVNA	CPAYARAFL	EEINAKKIPDNQV--KILKEIFRQFADACGQALD
TRG	PTQQLKLQGSVSVVNA	PLAYARAFL	DDINTKRYPDNKV--KILKEVFRQFVEACGQALA
KIAA1055	KLQLKLQGSVSVQVNAG	PLAYARAFL	DDINTKRYPDNKV--KILKEVFRQFVEACGQALA
CLASP-2	KLQLKLQGSVSVQVNAG	PLAYARAFL	DDINTKRYPDNKV--KILKEVFRQFVEACGQALA
CLASP-6	KLQLKLQGSVSVQVNAG	PLAYARAFL	DDINTKRYPDNKV--KILKEVFRQFVEACGQALA
CLASP-3	MLQMVLQGSVGTTVNQGP	LEVAQVF	LSE--IPSIDPKLFRHHNKIRLCFKDFTKRCEDALR
CLASP-4	QELQLKLQGCVSQVNAG	PLAYARAFL	INDSQASKYPPKKVSELKDMDRKFI--QACSIALE
CLASP-5	MLQMVLQGSVGATVNQGP	LEVAQVF	LAE--IPADPKLYRHHNKIRLCFKEFIMRCGEAVE
KIAA0716	PLTMCLNGVIDAAVNGGVS	RQEAFFV	KEYIILSHPEDGEKIAIRRELMLEQAQILEFGLA
DOCK2	PLSMLLNGLVDPAVMGG	FAKYEKAFFT	EYVRDHPEDQDKLTHIKDLIAWQIPFLGAGIK
DOCK3	LLSMLLNGLVIDAAVNGGI	ARYQEAFFD	DKOYINKHPGDAEKITOLKELMQEQVHVGLVGLA
DOCK180	PLSMLLNGLVDPAVMGG	FAKYEKAFFT	EYVRDHPEAHEKIEKLKDIAWQIPFLAEGIR
CONSENSUS	L M L+G V VN G	Y AFL + + P	II+ L
	L I	V V F	+ I

DOCK1=KIAA0716

DOCK3=KIAA0299

CLASPvariant=KIAA1055

ref 1.1

1 31
TCG ACT ATG AAT GCT GAT ACT GCT CCA ACA TCT TGT CCT TCC ATA TCT TCC CAG AAC
ser thr met asn ala asp thr ala pro thr ser pro cys pro ser ile ser ser gln asn

61 91
TCA AGC TCC TGC TCC AGC TTC CAG GAC CAG AAG ATC GCC AGC ATG TTC GAT CGG ACT TCC
ser ser ser cys ser ser phe gln asp gln lys ile ala ser met phe asp arg thr ser

121 151
AGA GTA CCC GCC AGC AGC ACT TCC TCA CCG GGG CTC CTC TTC ACA GAA CTG GCT GCT GCC
arg val pro ala ser ser thr ser ser pro gly leu leu phe thr glu leu ala ala ala

181 211
CTG GAT GCC GAA GGG GAA GGA ATC AGC GAA GTA CAA AGG AAA GCT GTC AGT GCA ATT CAC
leu asp ala glu gly glu gly ile ser glu val gln arg lys ala val ser ala ile his

241 271
AGC CTG CTA AGT TCT CAC GAC CTG GAC CCA CGC TGT GTC AAA CCA GAG GTG AAG GTC AAA
ser leu leu ser ser his asp leu asp pro arg cys val lys pro glu val lys val lys

301 311
ATC GCC GCC CTT TAC CTA CCT TTA GTT GGC ATC ATT TTG GAT GCT TTG CCA CAG CTC TGT
ile ala ala leu tyr leu pro leu val gly ile ile leu asp ala leu pro gln leu cys

361 391
GAC TTT ACA GTT GCA GAT ACT CGC AGA TAC CGC ACC AGT GGC TCG GAT GAA GAA CAA GAA
asp phe thr val ala asp thr arg arg tyr arg thr ser gly ser asp glu glu gln glu

421 451
GGA GCC GGT GCC ATT ACC CAG AAT GTG GCT CTG GCC ATA GCA GGG AAT AAT TTC AAT TTG
gly ala gly ala ile thr gln asn val ala leu ala ile ala gly asn asn phe asn leu

ref 2.1

461 511
AAA ACA AGT GGA ATA GTG CTG TCT TCC TTG CCC TAT AAG CAG TAC AAC ATG CTG AAC GCG
lys thr ser gly ile val leu ser ser leu pro tyr lys gln tyr asn met leu asn ala

541 571
GAC ACT ACT CGC AAC CTC ATG ATC TGC TTC CTC TGG ATC ATG AAA AAT GCT GAT CAG AGC
asp thr thr arg asn leu met ile cys phe leu trp ile met lys asn ala asp gln ser

601 631
CTC ATT AGG AAG TGG ATT GCT GAC CTG CCA TCA ACG CAG CTC AAC AGG ATT TTA GAT CTA
leu ile arg lys trp ile ala asp leu pro ser thr gln leu asn arg ile leu asp leu

661 691
CTT TTC ATC TGT GTG TTA TGT TTT GAG TAT AAG GGA AAA CAG AGT TCT GAC AAA GTC AGT
leu phe ile cys val leu cys phe glu tyr lys gly lys gln ser ser asp lys val ser

721 751
ACC CAA GTC CTG CAG AAG TCA AGG GAT GTC AAG GGC CGG CTG GAA GAG GCT TTG CTG CGT
thr gln val leu gln lys ser arg asp val lys ala arg leu glu glu ala leu leu arg

...
611

A

five
1 of 5

841 871
GGC CTA AAT GAA AAT TTG AGA TGG AAG AAA GAG CAG ACA CAT TGG CGG CAA GCT AAT GAG
gly leu asn glu asn leu arg trp lys lys glu gln thr his trp arg gln ala asn glu

901 931
AAG CTA GAT AAA ACA AAG GCC GAG TTA GAT CAA GAA GCC TTG ATC AGT GGC AAT CTG GCT
lys leu asp lys thr lys ala glu leu asp gln glu ala leu ile ser gly asn leu ala

961 991
ACA GAA GCA CAT TTA ATC ATC CTG GAT ATG CAG GAA AAC ATT ATC CAG GCG AGC TCG GCT
thr glu ala his leu ile ile leu asp met gln glu asn ile ile gln ala ser ser ala

1021 1051
CTG GAC TGT AAA GAC AGC CTG CTG GGA GGT GTT CTG AGG GTG CTG GTG AAT TCT CTG AAC
leu asp cys lys asp ser leu leu gly gly val leu arg val leu val asn ser leu asn

1081 1111
TGT GAT CAG AGT ACC ACC TAC CTG ACT CAC TGC TTT GCA ACA CTC CGT GCT CTC ATC GCC
cys asp gln ser thr thr tyr leu thr his cys phe ala thr leu arg ala leu ile ala

1141 1171
AAG TTT GGA GAC TTA CTC TTC GAA GAG GAG GTG GAA CAG TGT TTC GAC CTA TGT CAC CAA
lys phe gly asp leu leu phe glu glu val glu gln cys phe asp leu cys his gln

1201 1231
GTC CTG CAC CAC TGC AGC AGC ATG GAT GTC ACC CCG AGC CAA GCC TGT GCC ACC CTT
val leu his his cys ser ser ser met asp val thr arg ser gln ala cys ala thr leu

1261 1291
TAC CTC CTC ATG AGG TTC AGT TTT GGA GCC ACC AGT AAT TTT GCA AGA GTA AAG ATG CAA
tyr leu leu met arg phe ser phe gly ala thr ser asn phe ala arg val lys met gln

1321 1351
GTA ACC ATG TCC CTG GCA TCT TTG GTG GGA AGA GCA CCA GAC TTT AAT GAA GAG CAC CTG
val thr met ser leu ala ser leu val gly arg ala pro asp phe asn glu glu his leu

1381 1411
AGA AGA TCC TTG AGG ACA ATT TTG GCC TAT TCA GAA GAG GAC ACA GCC ATG CAG ATG ACT
arg arg ser leu arg thr ile leu ala tyr ser glu glu asp thr ala met gln met thr

1441 1471
CCT TTT CCC ACC CAG GTG GAG GAA CCT CTC TGT AAT CTG AAT AGC ATC TTA TAT GAC ACA
pro phe pro thr gln val glu glu leu leu cys asn leu asn ser ile leu tyr asp thr

1501 1531
GTG AAA ATG AGG GAA TTT CAG GAA GAT CCT GAG ATG CTT ATG GAT CTC ATG TAC AGA ATT
val lys met arg glu phe gln glu asp pro glu met leu met asp leu met tyr arg ile

1561 1591
GCC AAG AGT TAC CAG GCA TCT CCT GAT CTG CGG CTG ACC TGG CTC CAG AAC ATG GCA GAG
ala lys ser tyr gln ala ser pro asp leu arg leu thr trp leu gln asn met ala glu

1621 1651
AAA CAC ACC AAG AAG TGC TAC ACG GAG GCT GCC ATG TGC CTG GTG CAC GCC GCT GCG

1681 1711
TTA GTG GCT GAG TAT CTG AGC ATG CTG GAG GAC CAC AGC TAC CTG CCC GTG GGC AGT GTC
leu val ala glu tyr leu ser met leu glu asp his ser tyr leu pro val gly ser val

1741 1771
AGC TTC CAG AAT ATT TCT TCC AAT GTG CTG GAG GAG TCT GTG GTC TCT GAG GAC ACC CTG
ser phe gln asn ile ser ser asn val leu glu glu ser val val ser glu asp thr leu

1801 1831
TCA CCT GAC GAG GAT GGG GTG TGC GCA GGC CAG TAC TTC ACC GAG AGT GGC CTG GTA GGC
ser pro asp glu asp gly val cys ala gly gln tyr phe thr glu ser gly leu val gly

1861 1891
CTC CTG GAG CAG GCC GCG GAG CTC TTC AGC ACG GGA GGC TTA TAT GAG ACA GTT AAT GAG
leu leu glu gln ala ala glu leu phe ser thr gly gly leu tyr glu thr val asn glu

1921 1951
GTC TAC AAG CTG GTC ATC CCC ATC CTA GAA GCG CAT CGA GAA TTC CGG AAG CTG ACA CTC
val tyr lys leu val ile pro ile leu glu ala his arg glu phe arg lys leu thr leu

1981 2011
ACT CAC AGC AAG CTG CAG AGA GCC TTC GAC AGC ATC GTT AAC AAG GAT CAT AAG AGA ATG
thr his ser lys leu gln arg ala phe asp ser ile val asn lys asp his lys arg met

2041 2071
TTT GGA ACC TAC TTC CGA GTT GGT TTC TTT GGA TCC AAA TTT GGG GAT TTG GAT GAA CAG
phe gly thr tyr phe arg val gly phe phe gly ser lys phe gly asp leu asp glu gln

2101 2131
GAG TTT GTC TAC AAA GAG CCT GCA ATT ACC AAG CTT CCT GAG ATC TCA CAT AGA CTA GAG
glu phe val tyr lys glu pro ala ile thr lys leu pro glu ile ser his arg leu glu

2161 2191
GCA TTT TAT GGT CAA TGT TTT GGT GCA GAA TTT GTG GAA GTG ATT AAA GAC TCC ACT CCT
ala phe tyr gly gln cys phe gly ala glu phe val glu val ile lys asp ser thr pro
ref 4.1

2221 2251
↓
GTG GAC AAA ACC AAG TTG GAT CCT AAC AAG GCC TAC ATA CAG ATC ACT TTT GTG GAG CCC
val asp lys thr lys leu asp pro asn lys ala tyr ile gln ile thr phe val glu pro

2281 2311
TAG TTT GAT GAG TAT GAG ATG AAA GAC AGG GTC ACA TAC TTT GAG AAG AAT TTC AAC CTC
tyr phe asp glu tyr glu met lys asp arg val thr tyr phe glu lys asn phe asn leu

2341 2371
CGG AGG TTC ATG TAC ACC ACC CCG TTC ACC CTG GAG GGG CGG CCT CGG GGA GAG CTG CAT
arg arg phe met tyr thr pro phe thr leu glu gly arg pro arg gly glu leu his

2401 2431
GAG CAG TAC AGA AGG AAC ACA GTC CTG ACC ACT ATG CAC GCC TTC CCC TAC ATC AAG ACC
glu gln tyr arg arg asn thr val leu thr thr met his ala phe pro tyr ile lys thr

2461 2491
AGG ATC AGC GTC ATC CAG AAG GAG GAG TTT GTT TTG ACA CCG ATT GAA GTT GCC ATT GAA
arg ile ser val ile gln lys glu glu phe val leu thr pro ile glu val ala ile glu

2551
2521 GAC ATG AAG AAG AAG ACC CTG CAG TTA GCA GTT GCC ATT AAC CAG GAG CCG CCT GAT GCA
asp met lys lys thr leu gln leu ala val ala ile asn gln glu pro pro asp ala
2581 2611
AAG ATG CTT CAG ATG GTG CTG CAA GGC TCT GTG GGA GCT ACT GTA AAT CAG GGA CCA CTG
lys met leu gln met val leu gln gly ser val gly ala thr val asn gln gly pro leu
2641 2671
GAA GTA GCC CAA GTG TTT TTG GCT GAA ATT CCT GCT GAT CCA AAA CTC TAT CGA CAT CAC
glu val ala gln val phe leu ala glu ile pro ala asp pro lys leu tyr arg his his
2701 2731
AAC AAG TTG AGG TTA TGC TTT AAG GAA TTC ATC ATG AGA TGT GGT GAA GCT GTA GAG AAA
asn lys leu arg leu cys phe lys glu phe ile met arg cys gly glu ala val glu lys
2761 2791
AAC AAG CGT CTC ATC ACG GCA GAC CAG AGG GAA TAT CAG CAG GAA CTC AAA AAG AAC TAT
asn lys arg leu ile thr ala asp gln arg glu tyr gln gln glu leu lys asn tyr
2821 2851
AAC AAG CTA AAA GAG AAC CTC AGG CCA ATG ATC GAG CGG AAA ATT CCA GAA CTG TAC AAG
asn lys leu lys glu asn leu arg pro met ile glu arg lys ile pro glu leu tyr lys
2881 2911
CCA ATA TTC AGA GTT GAG AGT CAA AAG AGG GAC TCC CAC AGA TCT AGT TTC AGG AAA
pro ile phe arg val glu ser gln lys arg asp ser phe his arg ser ser phe arg lys
2941 2971
TGT GAA ACC CAG TTG TCA CAG GGC AGC TAA GAA AAG CCA TCT TCA TTC GTG GAG ACT GTG
cys glu thr gln leu ser gln gly ser OCH
ref 5.1
3001 3031
GCC CTG CAA CCC TGG AGA AGG ACT TGC TGG TAC TTA AAA AAT GGG ACA TTT GCC ACC CAG
3061 3091
GAC TGA CTG TAC ACT CCC TGA TCA GCC AGC ACT CTG GAA GCT TTG GGA TCC CAG GAA CCA
3121 3151
TGG AAT TAT TCC CAA ATG GAC TCT GAC CAG ATT TTT GCC ATA CTG GGG GGT GGC GGG ATG
3181 3211
GAG GAT GGG TAC TCA GGC ATG ACT GCG TAT TTA AAG TGT GTT TTT CCA CAA TGT ACC
3241 3271
AAA CAA GGC ATA AGC AGC TTC TCC TGC TGA CTG GCC AAT CAC TGC CCA TCT GAG AGA TGA
3301 3331
TTT CCT CTG GCC CAT ATT TGA ATT TAT TGG AGT AAC TCA AAT TGC CTG AGG AAA AAT GGA
3361 3391
AAA ATT ATC CAC CAG TCG ATT CAA ACT GAA TTT CAC TCT TTA TAG GAA GGC AGG GCA AAC
3421 3451
TTG TAG GAG TAC GAA ACA TTT TCA ATA AAT CTA CAA AGG GAA GCC TTA CTA CAA TTC CAA

AAA TCA TCA TGG TTG GAA ATT TGG GAG GAG ATT ATT TGT GAA CTT GTT ACC CTT TTG GTA
3541 3571
ATG GTG GAC TAA TTG CTG TAT AGT TAT TTT TGT TTT ATT ATT ACT GTT ACA TTA ATT TAA

3601 3631
CAT GCA TTT ATA GAA GAA TAC ATT CAA AGC ACT GAT GTA GGA GAT ACA CGG TAC TTG GAG

3661 3691
CAG TCA GCC AAA AAT CAC AGA TAC TGC TTT CAC TTA AAT GGA AAC AAT TCT CCG ATA ATG

3721 3751
CTT TGC TTT TTT TCT TAT GTC ACT CTT GTG TAC TAT CTA TTT TTC TCC TCT CTG GGA CCA

3781 3811
AGT TTC TTT TTA TAA AGC AAT AAT ATC TCT GTT TTC ATT TCA GAA CAT TGT GCT GTC TGT

3841 3871
CAG CAT ATG TAT ATC AGC TAC AAA ATA TAT TCA ACT TTG ACT TCT TTT GAC AAA GGA CTT

3901 3931
TAG GAA AAG GAG GAA CAA AGA CAT TAT TTG AGA ATT AAA TTA TAT ATT TTT AAT ATG ACT

3961 3991
GTG ACC TTG ACT GAT AAT AAA GAT GTA ATA AGA ATT GCA AGC TAA AAA AAA AAA AAA AAA

4021
AAC TCG

References

BAC sequences of Human CLASP 5

Ref 1.1

Sequence of BAC19 using primer HC5S11, which spans nucleotides 3-22 of the cDNA. Exon sequence is underlined and represents nucleotides 32-57.

CTCTCTGTCTTCATATCITCCAGGTTATAAAGNATTATTTACTAAAAGAANATTCANGCT
TATTCATTTAACTAGCTCAGTTAATCATGTATTCCTATAAAGGTTAGCTTATTAAATT
TTGACAAAACAATCAAACAATTCAAACCAGATCAAGTATGCTACCCTGAAGTTACACC
ACTAGCTAAGAATTAAACAATCTAAGTAATTGGTTCTCCCCAGGCTCAAGGCTCCCTGA
TCAGGTTAAGTAAAGCCAAGAATCCAATAAGCCCTATGAAATTAGAAACTCATAGAA
AAGTCTCAAATCTTCTTGTCTGACATTAGCCAATTGTTATATTATGCAAATAGAGGATT
NCAAGTAATAAGTTGGAACCTGTITACCAAGGTTTGCAAGCAGNCCTTAAGAGCTT
AACTGGTCATGCATTGAATGCCAGAGCAAAGAGGAATGGAGAGGGGNTGTAAGNGG
TTCCAATNTTACTGGAACCCACCACATCTTNGAAGTCTTGATACTTAACTGNGTGTA
GNCTCTTAGGCNTNTANTAANTAGAATCTATGGATTCTGTCTGTCNGCAAGNAG
TGCCTATGAAA

Ref 2.1

Sequence of BAC19 using primer HC5AS10b, which spans nucleotides 560-580 of the cDNA.
Exon sequence is underlined and represents nucleotides 510-553.

TGCGAGTAGTGTCCCGTTCAGCATGTTGACTGCTTATAGGGCTGAAGGGAGGCACG
ATTGGGGATGGAGGCCAGGAAGAAGTCAAGCAACAGAAAAATTGAGGCTTAACA
GTCAAGAACAGAAAATTCAAAGTGTCTTAAACATGACTGTACATCACTG
CTAGGCTGGAGATCTATTGCCAGTAGCCCTGCCTCCCTAGGCAGGGGAAGCTGTGTT
CTTGAGTAGCGTACTCAGCAAAGAGGCTCACCTGGGCAGTATTGAGCTAGGCTT
TCAGCCACCGTATCTGAGTACCTCTGCTTANGAGCAGTGTGCCCTGGTGTACACCCCT
GGGCCTTGTATCGTGTGCAATCCCAGTGATACAAAGAGGCTTCATGCTGCTAA
GATCTCCAAGTATTCTCCTCGTGTGGCAGCAGAGGGTTAGACTTNCAGGGAGA
AGGAAACTGGCTGGGTGCCATGAATAANCTGCTGTTCAAGANITAACCTCTTGTAC
ATAAGNGCAAAGGTATAACATAAAGGGNCATGAACTGCTCAACNAAATTNATCAAAT
CCATGTTGTGGAGTTCTTGTNATNGGAAGTTAACCCCTAA

Ref 3.1

Sequence of BAC13 using primer C5S3, which spans nucleotides 1086-1105 of the cDNA. Exon sequence is underlined and represents nucleotides 1110-1120.

CCCNCGCTCTTTGGCAANGTAANCNTGGGATGCTGTTCTTCCTCTTAATTAAAGAG
NAAGANTTTTTAGCTTCATACTTCTCTCTCAGGGGGACAAAAGTCACAGAGCATA
TTAAGTGGCANAACCCNAAGGTCTTAAGTCTTCTAGGAAGAAAGCAGATGCCCTGA
TTCTGTGGGAAGCCACCATGGAGAGGAAAGCAGTGGCTCCCATATTGAAGTGNGGA
CCTAACTCTAGAAAGTTAAAANGCCATTGCTGAAGGGCTATGACATGAGAACAGA
GATCAACTGAGTGACTTAGCAANTTCACTCTCTGTAAANACCTCTGGTGAGTGAGA
NTAAATCCTNTATGTGACGCCATTAGTCTTACAAANGTCATGCCNTAAATGCCAN
GAAGGNAGAAATGAATTCTACCGCCNGAGGAATGAGGATTATNCTGGGGGACA
TGCANAAATATTNNNCCCCNATTATTNATTATTGAGACNGAGTNTCGNT
CTAATCGCCCCCAGGCTGGNAGGTGGTCCCATCTNAANCNTANNTNGGA
AGNCCTCTTGNNGCCCCNGGGGGNGNAAAGNGAATTCCCTAAATGCCTNCANNCCC
CTCCCTGGGANTTATTGGGGNNTNTAAAGGCNGTGGCNG

Ref 4.1

Sequence of BAC13 using primer C5S7, which spans nucleotides 2196-2205 of the cDNA. . Exon sequence is underlined and represents nucleotides 2225-2231.

ACAAAAAACTAACCATCANTCTCTAAATCCCAACAANCTTTTAAGAATACCTAANG
AGCTCAACNAGGGGACTNTCCAANGCACTTAAATGCAGNCACNACNCCNNCAAG
AGNGGCAACTACTAATGGGCANATCTNAAAGAAAATAGNCAAAGGNNGGAATCA
TAATAGGAGGCAACCACCTAANGAACCAACTGGGCACTGGGCACTGCAACTGCAAC

TTCCTANNNTAGAGANGAGANAACGGGACATGGGAAGAGGNAAGCGAAGGGTTCA
AGGGGANGNAAGCGAGCAGANNCAGGGNCTCANACTNGNGGGNNNTGGGGGNTN
CTGNNNNCCCTACNCTTNGNANGAACAGNGNNGTTGANNTGGCTTNGANTA

Ref 5.1

Sequence of BAC13 using primer 122047F1, which spans nucleotides 3537-3556 of the cDNA.
Exon sequence is underlined and represents nucleotides 3000-3492. This region does not contain a
intron in this region

CCANNAGATTNTTGNAACGNNGTAGGCTTCCTTGTAGATTTATTGAAAATGTTCGT
ACTTCTACAAGTTGCCTGCCTTCTAAAGAGTGAAAANTCANTTGAATCGACTGG
TGGATAATTNTCCATTTTCCTCAGGCAATTNGAGTTACTCCAATAAATTCAAATAT
GGGCCAGAGGAATCATCTTCAGATGGGCAGTGATTGCCAGTCAGCAGGAGAAGC
TGCTTATGCCTTGTTGGTACATTGTGGAAAAACACACTTTAATAAATAACGCANTCATG
CCTGAGTACCCATCCTCCATCCCGCCACCCCCCAGTATGGCAAAATCTGGTCAGAGT
CCATTTGGGAATAATTCCATGGTCTGGGATCCCAAAGCTTCCAGAAGTGTGGCTG
ATCAANGGAGTGTACAGTCAGTCCTGGGTGGCAAAATGTCCCATTTTAAGTACCA
AGCAAAGGTTCCTTCTTCAAGGTTNCTAGGGCC

Figure

Multiple sequence alignment of Human CLASP proteins with internal exon borders indicated by a vertical line. Numbers in right margin correspond to References

HC2A	-----
KIAA	ASGNLDKNARFSAIYRQDSNKLSNDDMLKLLADFRKPEKMAKLPVILGNLDITIDNVSSD
rat	-----
HC4	-----
HC1	-----
HC3	-----
HC5	-----
HC2A	-----
KIAA	FPNYVNSSYIPTKQFETCSKTPITFEVEEFVPCIPKHTQPYTIYTNHLYVYPKYLKYDSQ
rat	-----
HC4	-----
HC1	-----
HC3	-----
HC5	-----
HC2A	-----VLHHHQNP EFYDEIK
KIAA	KSFAKARNIAICIEFKDSDEEDSQLKCIYGRPGGPVFRSAFAAVLHHHQNP EFYDEIK
rat	-----
HC4	-----
HC1	-----
HC3	-----
HC5	-----
HC2A	IELPTQLHEKHHLLLTFHVSCDNSSKGSTKKRDVVTQVGYSWLPLLKDGRRVTSEQHI
KIAA	IELPTQLHEKHHLLLTFHVSCDNSSKGSTKKRDVVTQVGYSWLPLLKDGRRVTSEQHI
rat	-----
HC4	-----
HC1	-----
HC3	-----
HC5	-----
HC2A	PVSANLPSGYLGQELGMGRHYGPEIKWVDGGKPLLKISTHLVSTVYTQDQHLHNFFQYC
KIAA	PVSANLPSGYLGQELGMGRHYGPEIKWVDGGKPLLKISTHLVSTVYTQDQHLHNFFQYC
rat	-----
HC4	-----
HC1	-----
HC3	-----GPGPARSTVSISLISNSARV
HC5	-----
HC2A	QKTESGAQALGNELVKYLKSLHAMEGHVMIAFLPTILNQLFRVLT-RATQEEVAVNVTRV
KIAA	QKTESGAQALGNELVKYLKSLHAMEGHVMIAFLPTILNQLFRVLT-RATQEEVAVNVTRV
rat	-----
HC4	-----MEIQVLIRFLSVILMQLFWVLPNMIHEDDVPISCPVM
HC1	-----MSFLPIILNQLFKVLV-QNEEDEITTTVTRV
HC3	-----NRSRSLNSNSNPDISGTPSPDDEVRSIIGSKGLDRSNSWVNTGGPKAAPWGNSPSPSAES
HC5	-----

B

Ref

HC2A |||HVVAVQ[RE]GLESHLRSYV[KYAYKAE[PYVASEYKTVHEELT[RE]TILKPSADFLTSN
 KIAA |||HVVAVQ[RE]GLESHLRSYV[KYAYKAE[PYVASEYKTVHEELT[RE]TILKPSADFLTSN
 rat
 HC4
 HC1
 HC3
 HC5

Cadherin
 Cleavage

KLLRYSWFFFDVLIKSMAQHLIENSKVKL[RNRQF]FPASYHAAETVVNMLMPHITQKFQD
 KLLKYSWFFFDVLIKSMAQHLIENSKVKL[RNRQF]FPASYHAAETVVNMLMPHITQKFQD

 KLLKYSWFFFEEIIAKSMATYLLEENKIKLTHGQRFPKAYHHALHSLFLAIT-IVESQYAE
 HVLKHSWFFFIAILKSMAQHLIDTNKIQLERPQRFPESYQNELDNLVMVLSDHVIWVKYKD
 SALQQAWFFFELMVKSMVHHLYFNDKLEAFRKSHFPERFMDDIAALVSTIASDIVSRFQK

HC2A NPEASKNANHSLAVFIKRCFTFMDRGFVFKQIN---NYIS--CFAPGD[KTLFEYKFEFL
 KIAA NPEASKNANHSLAVFIKRCFTFMDRGFVFKQIN---NYIS--CFAPGD[P]TLFEYKFEFL
 rat
 HC4
 HC1
 HC3
 HC5

IPKESRNVNYSLASFLKCCLTLMDRGFVFNLLIN---DYIS--GFSPKDPKVL[A]EYKFEFL
 ALEETRRATHSVARFLKRCFTFMDRGCVFKMVN---NYIS--MFSSGDLKTL[C]QYKFDFL
 DTEMVERLNTSLAFFLNDLLSVMDRGFVFSLIKSCYKQVSSKLYSLPNPSV[V]SLRDFL

RVVCNHEHYIPLNLP[RE]----PFGKGRIQR-----YQDLQL---DYSLTIEF
 RVVCNHEHYIPLNLP[RE]----PFGKGRIQR-----YQDLQL---DYSLTIEF

 QTICNHEHYIPLNLP[RE]----AFAKPKLQR-----VQDSNL---EYSLEI[E]
 QEVCQHEHFIPLCPIRSANIPDPLTPSES-----TOELHASDMPEYSVTNEF
 RIICSHEHYVTLNLP[RE]LTPPASPSPSVSSAT[S]QSSGFSTNVQDQKIANMFELS--VVF
 -----MNADTAPTSPCP[S]IS---SONSSCCSFQDQKIASMFDRTSRVPA

1.1

Cadherin
 EC motif

HC2A CRNHFLVG[G][RE]VGTA[LQEFRE]---VRLIAISV[RE]K[NLLIKH]SFDDRYASRSHQARIAT
 KIAA CRNHFLVG[G][RE]VGTA[LQEFRE]---VRLIAISV[RE]K[NLLIKH]SFDDRYASRSHQARIAT
 rat
 HC4
 HC1
 HC3
 HC5

CKHHFLVG[G][RE]TSIA[QDNYE]---IRYTAISV[RE]K[NLLIKH]AFDTRYQHKNQQAKIAQ
 CRKHFLIG[G][RE]VGFA[LQEDQD]---VRHLALAVLKNLMAKHSFDDRYREPRKQAOIAS
 RQQHYLAGLVLTELAVILDPAEGLFGLHKKVIMVWHNLLSSHSDPRYSDP[RE]I[KARVAM
 SSTS-SPG[G][RE]LAAALDAE[EGEGI]SEVQRKA[VSAIHSLLSSHDLDPRCVKPEVVKIAA

LYLPLFGLLIENVQRINVRDVSPFPVNAG-MTVKDES[LALPAVNPLVTPQKGSTLDNSLH
 LYLPLFGLLIENVQRINVRDVSPFPVNAG-MTVKDES[LALPAVNPLVTPQKGSTLDNSLH

 LYL[P]VGLLLENIQRLAGRDTLYSCAAMPNSASRDEFPCG----FTSP--AN--RG[RE]S
 LYMPLYGMLLDNMPRIYLKDLYPFTVNTSNGSRDDLSTNGGFQSQTAIKHANSVDT[RE]FS
 LYLPLIGIIMETVPOLYDFTETHNQRGRPI[RE]ATDDYESE----SG---SMIS
 LYLPLVGIILDALPQLCDFTVADTRRYR---TSGSDEEQE----GA---G[RE]IT

B

HC1A	KELLLGAISG I REFPYTTSTPNINSVRNADSRGS I STDGSNSLPERNSEKNSLD I HQQSS
K1AA	KELLLGAISG I ASP I PYTTSTPNINSVRNADSRGS I STDGSNSLPE I SEKNSLD I HQQSS
rat	
HC4	TPKPTAYGSFQNG-----HGIKREDSRGS I IP-EGATGFPDQGNTGEN----TQS
HC1	KDVLN I NIAFSS-----IAISTVNHADSRASLASLD I DSNPSTNEKSEKTNDCEKIPRPL
HC3	CTVAMAIAGT I SVPQ-----LTRPGS I LLT I TSGROHT-----
HC5	INVALAIAGNNFN-----LKTSG-IVLSSLPYKQYN-----

LPTs

2.

HC1A	TGNSVVRCOKLDDQSEIKSLLMCFLYILKSMSPDALFTYWN-KASTSELMDFFTISEVCL
KIAA	TGNSVVRCOKLDDQSEIKSLLMCFLYILKSMSPDALFTYWN-KASTSELMDFFTISEVCL
rat	
HC4	STRSSVSQYNRLDQEYIPLSLLMCYLIVKMISEDLLTYWN-KVSPQELINILILLEVCL
HC1	ALIGSTLRFPLRDQAETRSLLMCFLHIMKTISYETLIAYWQ-RAPSPEVSDFFSILVCL
HC3	-----TFSAELESSRLICLWLVLKNA-ADETVLLQKWFDTDLQLNRLLDLLYLCV
HCS	-----MLNADTTRNLMICFLWIMKN-ADQSLIRKWIADLPSTQLNRILDDLFLICV

HC2A	HQFQYMGKRYIARNQEGLG--PIVHDRKS-----	QTLPVSRNRTGMM
KIAA	HQFQYMGKRYIAR-----	TGMM
rat		
HC4	FHFRYMGKRN1ARVHDALSKHFGIDRKS-----	QTMPALRNRSGVMM
HC1	QNFRYLGKRNIIRKIAAAF--KFVQSTQNNGTLKGSNPSCQTSGLLAQWMHSTSRRHEGHM	
HC3	SCFEYKGKXVFERMNSLTFK--KSKDMDRAK-----	LEEA1ILGS1GARQEMV
HCS	LCFEYKGKQSSDKVSTCVLQ--KSRDVKAR-----	LEEALLRGEARGEMM

HC2A	HARLQQL-----GSLDNS-----LTFNHSGHSADVLHQSLLEANIATEVC
KIAA	HARLQQL-----GSLDNS-----LTFNHSGHSADVLHQSLLEANIATEVC
rat	-----
HC4	[ASPLQHL-----SSLESS-----FTLNHSSTTTEADI FHQALLEGNTATEVS
HC1	[HRSQTLPIRKG---NALSNPKL---LQMLDNTMTNSNEIDIVHHVDTTEANIATEGR
HC3	PRSRGQLEERSPSGSAFGSQENLPWKRKDMDTHWRQNTKEKLDKSRAEI EHEAL IDGNLATEAN
HCS	RRRAPGNDRFP----GLNEENLRWKKEQTHWRQANEKLDKTKAELDQEAL ISGNLATEAH

HC2A	LTALDTLSLFTLA FH NQ L LA D HGNPLMKKVFDVYLCFLQKHQSETALKNVFTALRS L IY
K1AA	LTAL T LSLFTLA F KNQ L LA D HGNPLMKKVFDVYLCFLQKHQSETALKNVFTALRS L IY
rat	-----KLSRGHSP L MKKVFDVYLCFLQKHQSE M ALKNVFTALRS L IY
HC4	LT V LETISFFTQCFC K THFLNN G HNP L MKKVFDIHLAF L KNGQSEV S LKHVFASLR A F I
HC1	LT I LDLVSLF T QTHQRC I Q C DC Q NSLMKRGFD T YMLFFQVNQS A TALKHV F ASLR E FC V
HC3	LI I LDTLEIVVQT V S--V T E S --KESI L G G V L K V L H SMAC Q NQS A VY L Q H CF A T Q R A L V S
HC5	LI I LD M QENIIQASS--ALDC--KD S L G G V L R V V N S L N C D Q S TT T YL H CF A T L R A IA

3.

HC2A	KFPSTFYEGRADMCAALCYEILKCCNSKLSSIRTEASQQLLYFLMRNNFDYTGKKSFVRTH
K1AA	KFPSTFYEGRADMCAALCYEILKCCNSKLSSIRTEASQQLLYFLMRNNFDYTGKKSFVRTH
rat	KFPSTFYEGRADMCAASLCYEVLKCCNSKLSSIRTEASQQLLYFLMRNNFDYTGKKSFVRTH
HC4	KFPSSAFFKGRVNMAAFCYEVLKCCNSKLSSIRTEASQQLLYFLMRNNFDYTGKKSFVRTH
HC1	KFPSSAFFQGPAADLCGSFCEYEVLKCCNHRSRSTQEASALLYFLMRQNFEFNQKQSIVRSH
HC3	MFPPELLFEEETEQCADLCLRLRLRHCSSSIGTIRSHPSASLYLLMRQNFEIIGN--NFARVK
HC5	KFGDLLFEEEVEQCFCDLCHQVLHHCCSSMDVTRSQACATLYLLMRFSFGATS--NFARVK

HC2A	LQVIISVSQLIADVVGIGETRFQQLSIIINNCANSDRLIKHTSFSSDVKDLTKRIRTVM
KIAA	LQVIISVSQLIADVVGIGGTRFQQLSIIINNCANSDRLIKHTSFSSDVKDLTKRIRTVM
rat	LQVIISLSQLIADVVGIGGTRFQQLSIIINNCANSDRLIKHTSFSSDVKDLTKRIRTVM
HC4	LQIIIAVSQQLIADVALSGGSRQESIIINNFANSDRPMLARAFPAEVVKDLTKRIRTVM
HC1	LQIIKAQSVQLIAD-AGIGGSRQHSLAITNNFANGDKQMNINFPAEVVKDLTKRIRTVM
HC3	MQVPMSISSLVGTSONENEEITRSIYKTEILTYEFDIIFIRETTFRDQYDNLVHNITM

Transmembrane

HC2A
KIAA
rat
HC4
HC1
HC3
HC5

ATAQMKEHENDPEMLVLDLQYS LAKSYASTPELRKTWLDMSARIHVKNGLSEAAMCYVHV
 ATAQMKEHENDPEMLVLDLQYS LAKSYASTPELRKTWLDMSARIHVKNGLSEAAMCYVHV
 ATAQMKEHENDPEMLVLDLQYS LAKSYASTPELRKTWLDMSARIHVKNGLSEAAMCYVHV
 ATAQMKEHEKDPEMLIDLQYS LAKSYASTPELRKTWLDMSAKIHVKNGDFSEAAMCYVHV
 ATAQMKEHEKDPEMLIDLQYS LANSYASTPELRRTWLESMAKIHARNGDLSEAAMCYIHI
 DTVKMKEHQEDPEMLIDLMYRIAKGYQTSPDRLRTWLQNMAKGHSERSHAEAAQCLVHS
 DTVKMREFOEDPEMLIDLMYRIAKSYQASPDLRLTWLQNMAEKHTKKCYTEAAMCLVHS

	domain	SH3	
HC2A	TALVAEYLTRKGV-	-	FRQGCTAFRVITPN
KIAA	TALVAEYLTRKEA-	-VQWEPPPLPHSHSACLRSRGGVFRQGCTAFRVITPN	
rat	TALVAEYLTRKEAD-	-LALQREPPVFVPSHTSCQRKSRRGMFRQGCTAFRVITPN	
HC4	AALVAEFLHRKKL-	-	FPGNGCSAFKKITPN
HC1	AALIAEYLKRKGYWKVEKIC	TASLLSEDT	
HC3	AALVAEYLSMLED-	PCDSNSLLTPSGGSMSFMGWPFLSITPN	RKYPVGCVTFQNISSN
HC5	AALVAEYLSMLED-	-	HSYLPVGSVSFQNISSN

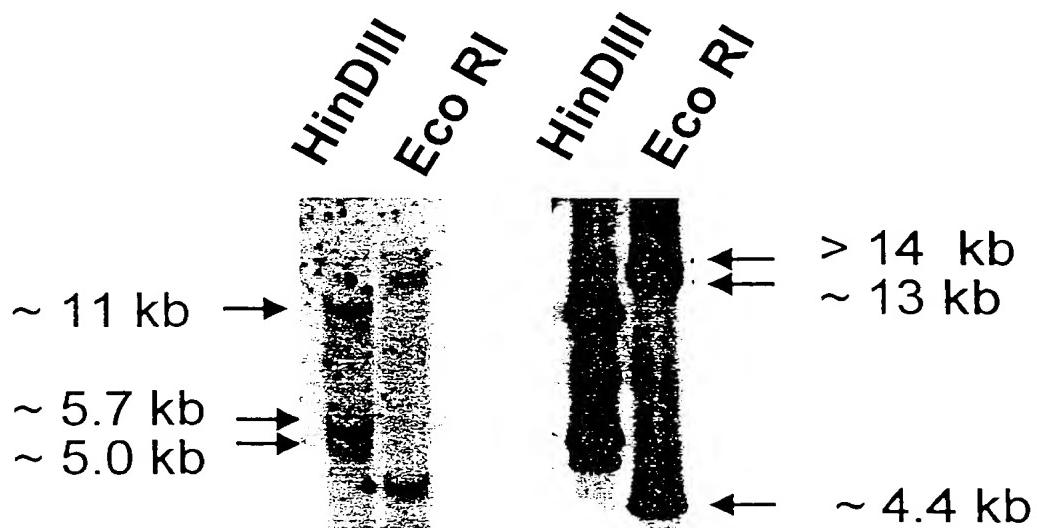
	ITAM
HC2A	I DEEASM MEDV GMQD-----VHFNE DVL MELLE Q CAD GLW KA ERYELIADI YKL I I PI
KIAA	I DEEASM MEDV GMQD-----VHFNE DVL MELLE Q CAD GLW KA ERYELIADI YKL I I PI
rat	I DEEASM MEDV GMQD-----VHFNE DVL MELLE Q CAD GLW KA ERL RAG LLTS INSSSP
HC4	I DEEGAM KEDAG MMD-----VHYSEE VLL ELC Q CVN GLW KA ERYETI SE ISKLIGPI
HC1	I KE GA EKA DSG GMHD-----TPYNF N I LVE QLY MC GE F LW KSE RYELIADVNKP I I AV
HC3	VLEES A VSD DVV S PDEEG IC SG KY FT E S GLV GL LEQ AA AS F SMAG MYE A V NE V K V L I I PI
HC5	VLEES VV S E D T L S P D E D G V C A G Q Y FT E S GLV GL LEQ AA EL F ST GG LY ET V NE V K V L I I PI

	ITAM	ITAM	ITAM	ITAM
HC2A	YEKPRD-----			
KIAA	YEKRRD FERLA HLY DT I H RAY SKV TE VMH SG RRL LG I F R WAFF G QAA QY QFT DSET DVE			
rat	SMKSGGT LETT HLY DT I H RPH SKV TE V ITR-----A-----AGSW DLL PG GL FG Q			
HC4	YENRREFEN LT QVY RT I H GAY T K I L E V M H T K R L L G-----T FFR V A FY G Q			
HC1	FEK QRD FK L S D I Y D I H R S Y L K V A E V V N S E K R L F G-----H YY R V A FY G Q			
HC3	HEANRDAKKL S T I H G K L Q E A F S K I V H Q S T G W E R M F G-----T Y F R V G F Y G-----			
HC5	LEAH R E F R K L T L H S K L Q R A F D S I V N K D H -----K R M F G-----T Y F R V G F F G-----			

	ITAM	ITAM
HC2A	- F F E D E D G K E Y I Y K E P K L T P L S E I S Q R L L K I Y S D R F G S E N V K M I Q D S G K V N P K D L D S H Y A	
KIAA	G F F E D E D G K E Y I Y K E P K L T P L S E I S Q R L L K I Y S D R F G S E N V K M I Q D S G K V N P K D L D S H Y A	
rat	G F F E D E D G K E Y I Y K E P K L T P L S E I S Q R L L K I Y S D R F G S E N V K M I Q D S G K V N P K D L D S H Y A	
HC4	S F F E E D G K E Y I Y K E P K L T G L S E I S L R L V K I Y G E K F G T E N V K I I Q D S D K V N A K E L D P H Y A	
HC1	G F F E E E G K E Y I Y K E P K L T G L S E I S Q R L L K I Y A D R F G A D M V K I I Q D S N K V N P K D L D P H Y A	
HC3	T K F G D L D E Q F V Y K E P A I T K L A E I S H R I E G Y G E R F G E D V V E V I K D S N P V D K C K L D P N K A	
HC5	S K F G D L D E Q F V Y K E P A I T K L P E I S H R I E G Y G O C F G A E F V E V I K D S T P V D K T K L D P N K A	

4.1

	ITAM
HC2A	Y I D V T H V I P F F D E K E L Q E R K T E F E R S H N I R R F M F E M P F T Q T G K R Q G G V E E Q C K R R T I L T A
KIAA	Y I D V T H V I P F F D E K E L Q E R K T E F E R S H N I R R F M F E M P F T Q T G K R Q G G V E E Q C K R R T I L T A
rat	Y I D V T H V T P F F D E K E L Q E R K T E F E R S H N I R R F M F E M P F T Q T G K R Q G G V E E Q C K R R T I L T A
HC4	H I D V T V V K P Y F D D O K E L T E R K T E F E R S H N I S R F V F E A P Y T I L S G K K Q G C I E E Q C K R R T I L T T
HC1	Y I Q V T V V T P F F E E K E I E D R K T D F E M M H H I N R F V F E T P F T I L S G K K H G G V A E Q C K R R T I L T T
HC3	Y I Q V T V V T P F F E E K E I E D R K T D F E M M H H I N R F V F E T P F T I L S G K K H G G V A E Q C K R R T I L T T



-111

CGGTAACCGCCATTTGTCTCCTGTAACAATTACGGCCGTGTAACTGTGAATCTTCAAAGCCTCAGTTTATGACC
CTGTGGAGGCCAGTGGACTTGAAAGGACTTCTG -1

1/1

31/11

ATG ACA CAC CTG AAC AGC CTG GAT GTG CAG CTT GCC CAG GAG CTC GGG GAC TTC ACT GAT
Met thr his leu asn ser leu asp val gln leu ala gln glu leu gly asp phe thr asp
61/21 91/31

GAC GAC TTG GAC GTG GTG TTC ACG CCA AAG GAA TGT AGG ACT TTG CAG CCC TCT TTG CCG
asp asp leu asp val val phe thr pro lys glu cys arg thr leu gln pro ser leu pro
121/41 151/51

GAG GAA GGG GTT GAA CTG GAC CCT CAT GTC AGG GAC TGT GTT CAG ACC TAC ATC CGT GAG
glu glu gly val glu leu asp pro his val arg asp cys val gln thr tyr ile arg glu
181/61 211/71

TGG CTA ATC GTG AAC CGG AAA AAC CAA GGA AGT CCA TAA ATC TGT GGC TTT AAA AAG ACT
trp leu ile val asn arg lys asn gln gly ser pro glu ile cys gly phe lys lys thr
241/81 271/91

GGA TCT CGA AAA GAT TTT CAC AAG ACG CTT CCG AAA CAG ACG TTT GAG TCG GAA ACC TTG
gly ser arg lys asp phe his lys thr leu pro lys gln thr phe glu ser glu thr leu
301/101 331/111

GAG TGC AGT GAA CCC GCT GCT CAG GCA GGC CCC CGC CAC TTA AAC GTG CTG TGC GAC GTG
glu cys ser glu pro ala ala gln ala gly pro arg his leu asn val leu cys asp val
361/121 391/131

TCT GGG AAA GGC CCC GTC ACT GCC TGT GAC TTT GAC CTC CGC AGC CTG CAG CCT GAC AAG
ser gly lys gly pro val thr ala cys asp phe asp leu arg ser leu gln pro asp lys
421/141 451/151

CGG CTA GAA AAC CTC CTG CAG CAA GTG AGT GCC GAG SAC TTT GAG AAG CAG AAC GAG GAG
arg leu glu asn leu leu gln gln val ser ala glu asp phe glu lys gln asn glu glu
481/161 511/171

GCC CGG AGG ACC AAC AGG CAG GCC GAG CTC TTT GCC ATT TAC CCA TCA GTG GAC GAG GAG
ala arg arg thr asn arg gln ala glu leu phe ala leu tyr pro ser val asp glu glu
541/181 571/191

GAT GCT GTG GAA ATA CGT CCA GTA CCA GAA TGT CCG AAG GAA CAC CTG GGC AAC AGA ATA
asp ala val glu ile arg pro val pro glu cys pro lys glu his leu gly asn arg ile
601/201 631/211

TTG GTC AAG TTG CTG ACC TTG AAG TTC GAG ATT GAA ATT GAG CCC CTG TTT GCC AGC ATT
leu val lys leu leu thr leu lys phe glu ile gln ile glu pro leu phe ala ser ile
661/221 691/231

GCC CTC TAC GAT GTT AAA GAA AGG AAA AAG ATC TCA GAA AAT TTT CAC TGT GAC CTG AAC
ala leu tyr asp val lys glu arg lys lys ile ser glu asn phe his cys asp leu asn
721/241 751/251

TCT GAC CAG TTC AAA GGA TTT CTG CGA GCT CAC ACG CCT TCA GTG GCC GCA TCA AGT CAG
ser asp gln phe lys gly phe leu arg ala his thr pro ser val ala ala ser ser gln
781/261 811/271

GCG AGA TCT GCA GTC TTC TCA GTC ACC TAC CCG TCC TCA GAC ATC TAC CTG GTA GTC AAG
ala arg ser ala val phe ser val thr tyr pro ser ser asp ile tyr leu val val lys
841/281 871/291

ATT GAA AAA GTC CTG CAG CAG GGA GAT ATT GGA SAC TGT GCA GAG CCC TAC ACG GTT ATC
ile glu lys val leu gln gln gly asp ile gly asp cys ala glu pro tyr thr val ile
901/301 931/311

AAA GAA AGT GAT GGT GGA AAG AGT AAA GAA AAG ATT GAA AAA CTA AAA CTC CAA GCT GAA
lys glu ser asp gly gly lys ser lys glu lys ile glu lys leu lys leu gln ala gln

1021/341

TCA AGC TTC TTC AAT GTC TCC ACC CTT GAG AGG GAG GTA ACT GAT GTG GAC TCT GTG GTT
ser ser phe asn val ser thr leu glu arg glu val thr asp val asp ser val val
1081/361 1111/371

GGG AGA AGC CCA GTG GGT GAA CGG AGG ACA TTG GCC CAA TCT AGA AGG CTT TCT GAA AGA
gly arg ser pro val gly glu arg arg thr leu ala gln ser arg arg leu ser glu arg
1141/381 1171/391

GCC CTC TCC TTG GAG GAA AAT GGG GTT GGA TCC AAC TTC AAA ACC TCC ACT CTG AGC GTT
ala leu ser leu glu glu asn gly val gly ser asn phe lys thr ser thr leu ser val
1201/401 1231/411

AGC AGC TTT TTC AAG CAG GAA GGA GAT CGC CTT AGC GAT GAA GAC TTA TTC AAG TTT TTA
ser ser phe phe lys gln glu gly asp arg leu ser asp glu asp leu phe lys phe leu
1261/421 1291/431

GCT GAC TAC AAA AGA TCA TCA TCC TTA CAG AGA CGA GTC AAG TCA ATT CCA GGC TTG CTA
ala asp tyr lys arg ser ser leu gln arg arg val lys ser ile pro gly leu leu
1321/441 1351/451

AGA CTG GAG ATT TCT ACA GCT CCA GAG ATC ATC AAT TGC TGT CTG ACT CCT GAA ATG CTG
arg leu glu ile ser thr ala pro glu ile ile asn cys cys leu thr pro glu met leu
1381/461 1411/471

CCC GTG AAA CCC TTT CCT GAA AAC CGG ACA CGC CCG CAC AAA GAG ATT TTG GAA TTT CCA
pro val lys pro phe pro glu asn arg thr arg pro his lys glu ile leu glu phe pro
1441/481 1471/491

ACA CGA GAA GTA TAT GTC CCT CAC ACT GTG TAC AGA AAC CTT CTC TAT GTC TAC CCA CAG
thr arg glu val tyr val pro his thr val tyr arg asn leu leu tyr val tyr pro gln
1501/501 1531/511

AGG CTG AAC TTT GTA AAC AAA CTA GCA TCA GCC CGG AAC ATT ACA ATA AAG ATC CAG TTT
arg leu asn phe val asn lys leu ala ser ala arg asn ile thr ile lys ile gln phe
1561/521 1591/531

ATG TGT GGA GAA GAT GCT AGC AAT GCG ATG CCG GTC ATC TTT GGA AAA TCC AGC GGG CCT
met cys gly glu asp ala ser asn ala met pro val ile phe gly lys ser ser gly pro
1621/541 1651/551

GAA TTT CTG CAG GAA GTG TAC ACA GCT GTT ACA TAC CAT AAT AAG TCT CCT GAC TTT TAT
glu phe leu gln glu val tyr thr ala val thr tyr his asn lys ser pro asp phe tyr
1681/561 1711/571

GAA GAA GTG AAA ATT AAG CTC CCC GCT AAG CTC ACA GTA AAT CAC CAC CTC CTG TTC ACC
glu glu val lys ile lys leu pro ala lys leu thr val asn his his leu leu phe thr
1741/581 1771/591

TTC TAC CAT ATC AGC TGT CAG CAG AAG CAA GGA GCC TCC GTG GAA ACT CTC CTG GGA TAT
phe tyr his ile ser cys gln gln lys gln gly ala ser val glu thr leu leu gly tyr
1801/601 1831/611

TCA TGG CTG CCA ATT CTC TTA AAT GAA CGT CTT CAA ACT GGA TCC TAC TGT CTC CCA GTT
ser trp leu pro ile leu leu asn glu arg leu gln thr gly ser tyr cys leu pro val
1861/621 1891/631

GCC TTG GAA AAA TTG CCA CCC AAC TAC TCC ATG CAT TCT GCT GAG AAA GTC CCA TTA CAG
ala leu glu lys leu pro pro asn tyr ser met his ser ala glu lys val pro leu gln
1921/641 1951/651

AAT CCT CCC ATT AAG TGG GCT GAA GGA CAT AAG GGA GTA TTT AAT ATT GAA GTG CAA GCT
asn pro pro ile lys trp ala glu gly his lys gly val phe asn ile glu val gln ala
1981/661 2011/671

GTT TCT TGT CAA CAC ACC CAG GAC AAC CAC CTG GAG AAG TTC TTC ACC CTC TGC CAC TCC
val ser ser val his thr gln asp asn his leu glu lys phe phe thr leu cys his ser
2041/681 2071/691

CTG GAG AGC CAG GTG ACC TTC CCC ATC CGC GTG CTG GAT CAG AAA ATC AGC GAG ATG GCG
leu glu ser thr val phe pro ile arg val leu asn gln lys ile ser glu met ala

2161/721

GTG CTC TTC CTG CAC CTG GTG CTG GAC AAG CTC TTC CAG CTG TCC GTG CAG CCC ATG GTC
val leu phe leu his leu val leu asp lys leu phe gln leu ser val gln pro met val
2221/741

ATC GCT GGC CAG ACA GCC AAC TTC TCC CAG TTT GCC TTC GAG TCC GTG GTG GCC ATC GCC
ile ala gly gln thr ala asn phe ser gln phe ala phe glu ser val val ala ile ala
2281/761

AAC AGT CTG CAC AAC AGC AAG GAC CTG AGC AAG GAC CAG CAT GGG AGG AAC TGC CTG CTG
asn ser leu his asn ser lys asp leu ser lys asp gln his gly arg asn cys leu leu
2341/781

GCT TCC TAC GTG CAC TAC GTC TTC CGC CTG CCA GAG GTG CAA AGG GAT GTG CCC AAG TCA
ala ser tyr val his tyr val phe arg leu pro glu val gln arg asp val pro lys ser
2401/801

GGC GCT CCC ACT GCC CTC CTA GAC CCT CGG AGC TAC CAC ACG TAT GGC CGC ACA TCA GCT
gly ala pro thr ala leu leu asp pro arg ser tyr his thr tyr gly arg thr ser ala
2461/821

GCT GCT GTG AGT TCA AAG CTG CTG CAG GCC CGG GTG ATG AGC AGC AGT AAC CCA GAC CTC
ala ala val ser ser lys leu leu gln ala arg val met ser ser ser asn pro asp leu
2521/841

GCG GGG ACA CAC TCC GCA GCA GAC GAG GAA GTG AAG AAC ATC ATG TCT TCA AAG ATC GCC
ala gly thr his ser ala ala asp glu glu val lys asn ile met ser ser lys ile ala
2581/861

GAT CGC AAC TGC AGC CGA ATG TCT TAC TAT TGC TCT GGC AGT AGT GAT GCT CCA AGT TCA
asp arg asn cys ser arg met ser tyr tyr cys ser gly ser ser asp ala pro ser ser
2641/881

CCT GCA GCC CCA AGG CCA GCC AGC AAA AAG CAT TTC CAT GAG GAG CTT GCC CTT CAG ATG
pro ala ala pro arg pro ala ser lys lys his phe his glu glu leu ala leu gln met
2701/901

GTG GTC AGC ACC GGA ATG GTG AAA AGC ATG GCC CAG CAC GTA CAT AAC ATG GAC AAA CGG
val val ser thr gly met val lys ser met ala gln his val his asn met asp lys arg
2761/921

GAC AGT TTT CGG AGG ACT CGT TTT TCT GAC TGT TTC ATG GAT GAC ATA ACT ACT ATT GTT
asp ser phe arg arg thr arg phe ser asp arg phe met asp asp ile thr thr ile val
2821/941

AAT GTG GTC ACC TCG GAA ATT GCA GCC CTT TTA GTA AAA CCA CAG AAG GAA AAT GAA CAG
asn val val thr ser glu ile ala ala leu leu val lys pro gln lys glu asn glu gln
2881/961

GCG GAA AAG ATG AAC ATC AGC CTG GCT TTC TTC TTG TAT GAC CTT CTC TCC CTC ATG GAT
ala glu lys met asn ile ser leu ala phe phe leu tyr asp leu leu ser leu met asp
2941/981

CGG GGC TTT GTG TTT AAC CTC ATC AGA CAT TAT TGC AGC CAG CTG TCA GCC AAG CTC AGT
arg gly phe val phe asn leu ile arg his tyr cys ser gln leu ser ala lys leu ser
3001/1001

AAC CTT CCA ACG CTC ATT TCC ATG AGG CTA GAG TTC CTG AGA ATC CTC TGT AGC CAT GAG
asn leu pro thr leu ile ser met arg leu glu phe leu arg ile leu cys ser his glu
3061/1021

CAT TAC CTC AAT CTG AAC CTT TTT ATG AAT GCT GAT ACT GCT CCA ACA TCT CCT TGT
his tyr leu asn leu asn leu phe phe met asn ala asp thr ala pro thr ser pro cys
3121/1041

CCT TCC ATA TCT TCC CAG AAC TCA AGC TCC TGC TCC AGC TTC CAG GAC CAG AAG ATC GCC
pro ser ile ser ser gln asn ser ser ser cys ser ser phe gln asp gln lys ile ala
3181/1061

AGC ATG TTC GAT CTG ACT TCC GAG TAC CGC CAG CAG CAC TTC ACC GGG CTC CTC TTC
glu asp val leu thr gln gln his phe leu thr gly leu leu phe

2191/731

GTG CTC TTC CTG CAC CTG GTG CTG GAC AAG CTC TTC CAG CTG TCC GTG CAG CCC ATG GTC
leu phe gln leu ser val gln pro met val
2251/751

ATC GCT GGC CAG ACA GCC AAC TTC TCC CAG TTT GCC TTC GAG TCC GTG GTG GCC ATC GCC
ile ala gly gln thr ala asn phe ser gln phe ala phe glu ser val val ala ile ala
2311/771

AAC AGT CTG CAC AAC AGC AAG GAC CTG AGC AAG GAC CAG CAT GGG AGG AAC TGC CTG CTG
asn ser leu his asn ser lys asp leu ser lys asp gln his gly arg asn cys leu leu
2371/791

GCT TCC TAC GTG CAC TAC GTC TTC CGC CTG CCA GAG GTG CAA AGG GAT GTG CCC AAG TCA
ala ser tyr val his tyr val phe arg leu pro glu val gln arg asp val pro lys ser
2431/811

GGC GCT CCC ACT GCC CTC CTA GAC CCT CGG AGC TAC CAC ACG TAT GGC CGC ACA TCA GCT
gly ala pro thr ala leu leu asp pro arg ser tyr his thr tyr gly arg thr ser ala
2491/831

GCT GCT GTG AGT TCA AAG CTG CTG CAG GCC CGG GTG ATG AGC AGC AGT AAC CCA GAC CTC
ala ala val ser ser lys leu leu gln ala arg val met ser ser ser asn pro asp leu
2551/851

GCG GGG ACA CAC TCC GCA GCA GAC GAG GAA GTG AAG AAC ATC ATG TCT TCA AAG ATC GCC
ala gly thr his ser ala ala asp glu glu val lys asn ile met ser ser lys ile ala
2611/871

GAT CGC AAC TGC AGC CGA ATG TCT TAC TAT TGC TCT GGC AGT AGT GAT GCT CCA AGT TCA
asp arg asn cys ser arg met ser tyr tyr cys ser gly ser ser asp ala pro ser ser
2671/891

CCT GCA GCC CCA AGG CCA GCC AGC AAA AAG CAT TTC CAT GAG GAG CTT GCC CTT CAG ATG
pro ala ala pro arg pro ala ser lys lys his phe his glu glu leu ala leu gln met
2731/911

GTG GTC AGC ACC GGA ATG GTG AAA AGC ATG GCC CAG CAC GTA CAT AAC ATG GAC AAA CGG
val val ser thr gly met val lys ser met ala gln his val his asn met asp lys arg
2791/931

GAC AGT TTT CGG AGG ACT CGT TTT TCT GAC TGT TTC ATG GAT GAC ATA ACT ACT ATT GTT
asp ser phe arg arg thr arg phe ser asp arg phe met asp asp ile thr thr ile val
2851/951

AAT GTG GTC ACC TCG GAA ATT GCA GCC CTT TTA GTA AAA CCA CAG AAG GAA AAT GAA CAG
asn val val thr ser glu ile ala ala leu leu val lys pro gln lys glu asn glu gln
2911/971

GCG GAA AAG ATG AAC ATC AGC CTG GCT TTC TTG TAT GAC CTT CTC TCC CTC ATG GAT
ala glu lys met asn ile ser leu ala phe phe leu tyr asp leu leu ser leu met asp
2971/991

CGG GGC TTT GTG TTT AAC CTC ATC AGA CAT TAT TGC AGC CAG CTG TCA GCC AAG CTC AGT
arg gly phe val phe asn leu ile arg his tyr cys ser gln leu ser ala lys leu ser
3031/1011

AAC CTT CCA ACG CTC ATT TCC ATG AGG CTA GAG TTC CTG AGA ATC CTC TGT AGC CAT GAG
asn leu pro thr leu ile ser met arg leu glu phe leu arg ile leu cys ser his glu
3091/1031

CAT TAC CTC AAT CTG AAC CTT TTT ATG AAT GCT GAT ACT GCT CCA ACA TCT CCT TGT
his tyr leu asn leu asn leu phe phe met asn ala asp thr ala pro thr ser pro cys
3151/1051

CCT TCC ATA TCT TCC CAG AAC TCA AGC TCC TGC TCC AGC TTC CAG GAC CAG AAG ATC GCC
pro ser ile ser ser gln asn ser ser ser cys ser ser phe gln asp gln lys ile ala
3211/1071

3301/1101 3331/1111
GCT GTC AGT GCA ATT CAC AGC CTG CTA AGT TCT CAC GAC CTG GAC CCA CGC TGT GTC AAA
ala val ser ala ile his ser leu leu ser ser his asp leu asp pro arg cys val lys
3361/1121 3391/1131
CCA GAG GTG AAG GTC AAA ATC GCC GCC CTT TAC CTA CCT TTA GTT GGC ATC ATT TTG GAT
pro glu val lys val lys ile ala ala leu tyr leu pro leu val gly ile ile leu asp
3421/1141 3451/1151
GCT TTG CCA CAG CTC TGT GAC TTT ACA GTT GCA GAT ACT CGC AGA TAC CGC ACC AGT GGC
ala leu pro gln leu cys asp phe thr val ala asp thr arg arg tyr arg thr ser gly
3481/1161 3511/1171
TCG GAT GAA CAA GAA GGA GCC GGT GCC ATT AAC CAG AAT GTG GCT CTG GCC ATA GCA
ser asp glu glu gln glu gly ala gly ala ile asn gln asn val ala leu ala ile ala
3541/1181 3571/1191
GGG AAT AAT TTC AAT TTG AAA ACA AGT GGA ATA GTG CTG TCT TCC TTG CCC TAT AAG CAG
gly asn asn phe asn leu lys thr ser gly ile val leu ser ser leu pro tyr lys gln
3601/1201 3631/1211
TAC AAC ATG CTG AAC GCG GAC ACT ACT CGC AAC CTC ATG ATC TGC TTC CTC TGG ATC ATG
tyr asn met leu asn ala asp thr thr arg asn leu met ile cys phe leu trp ile met
3661/1221 3691/1231
AAA AAT GCT GAT CAG AGC CTC ATT AGG AAG TGG ATT GCT GAC CTG CCA TCA ACG CAG CTC
lys asn ala asp gln ser leu ile arg lys trp ile ala asp leu pro ser thr gln leu
3721/1241 3751/1251
AAC AGG ATT TTA GAT CTA CTT TTC ATC TGT GTG TTA TGT TTT GAG TAT AAG GGA AAA CAG
asn arg ile leu asp leu leu phe ile cys val leu cys phe glu tyr lys gly lys gln
3781/1261 3811/1271
AGT TCT GAC AAA GTC AGT ACC CAA GTC CTG CAG AAG TCA AGG GAT GTC AAG GCC CGG CTG
ser ser asp lys val ser thr gln val leu gln lys ser arg asp val lys ala arg leu
3841/1281 3871/1291
GAA GAG GCT TTG CTG CGT GGG GAA GGG GCC AGA GGG GAG ATG ATG CGC CGC CGG GCT CCA
glu glu ala leu leu arg gly glu gly ala arg gly glu met met arg arg arg ala pro
3901/1301 3931/1311
GGG AAC GAC CGA TTT CCA GGC CTA AAT GAA AAT TTG AGA TGG AAG AAA GAG CAG ACA CAT
gly asn asp arg phe pro gly leu asn glu asn leu arg trp lys lys glu gln thr his
3961/1321 3991/1331
TGG CGG CAA GCT AAT GAG AAG CTA GAT AAA ACA AAG GCC GAG TTA GAT CAA GAA GCC TTG
trp arg gln ala asn glu lys leu asp lys thr lys ala glu leu asp gln glu ala leu
4021/1341 4051/1351
ATC AGT GGC AAT CTG GCT ACA GAA GCA CAT TTA ATC ATC CTG GAT ATG CAG GAA AAC ATT
ile ser gly asn leu ala thr glu ala his leu ile ile leu asp met gln glu asn ile
4081/1361 4111/1371
ATC CAG GCG AGC TCG GCT CTG GAC TGT AAA GAC AGC CTG CTG GGA GGT GTT CTG AGG GTG
ile gln ala ser ser ala leu asp cys lys asp ser leu leu gly gly val leu arg val
4141/1381 4171/1391
CTG GTG AAT TCT CTG AAC TGT GAT CAG AGT ACC ACC TAC CTG ACT CAC TGC TTT GCA ACA
leu val asn ser leu asn cys asp gln ser thr thr tyr leu thr his cys phe ala thr
4201/1401 4231/1411
CTG CGT GCT CTC ATC GCC AAG TTT GGA GAC TTA CTG TTC GAA GAG GAG GTG GAA CAG TGT
leu arg aia leu ile ala lys phe gly asp leu leu phe glu glu glu val glu gln cys
4261/1421 4291/1431
TTC GAC CTA TGT CAC CAA GTC CTG CAC CAC TGC AGC AGC AGC ATG GAT GTC ACC CGG AGC
phe asp leu cys his gln val leu his his cys ser ser ser met asp val thr arg ser
4321/1441 4351/1451
CAA GCC TGT GCC ACC CTT TAC CTC CTC ATG AGG TTC AGT TTT GGA GCC ACC AGT AAT TTT
gln ala cys ala thr leu tyr leu met arg phe ser phe gly ala thr ser asn phe

4441/1481

TTT AAT GAA GAG CAC CTG AGA AGA TCC TTG AGG ACA ATT TTG GCC TAT TCA GAA GAG GAC
phe asn glu glu his leu arg arg ser leu arg thr ile leu ala tyr ser glu glu asp
4501/1501 4531/1511
ACA GCC ATG CAG ATG ACT CCT TTT CCC ACC CAG GTG GAG GAA CTT CTC TGT AAT CTG AAT
thr ala met gln met thr pro phe pro thr gln val glu glu leu leu cys asn leu asn
4561/1521 4591/1531
AGC ATC TTA TAT GAC ACA GTG AAA ATG AGG GAA TTT CAG GAA GAT CCT GAG ATG CTT ATG
ser ile leu tyr asp thr val lys met arg glu phe gln glu asp pro glu met leu met
4621/1541 4651/1551
GAT CTC ATG TAC AGA ATT GCC AAG AGT TAC CAG GCA TCT CCT GAT CTG CGG CTG ACC TGG
asp leu met tyr arg ile ala lys ser tyr gln ala ser pro asp leu arg leu thr trp
4681/1561 4711/1571
CTC CAG AAC ATG GCA GAG AAA CAC ACC AAG AAG AAG TGC TAC ACG GAG GCT GCC ATG TGC
leu gln asn met ala glu lys his thr lys lys cys tyr thr glu ala ala met cys
4741/1581 4771/1591
CTG GTG CAC GCC GCT GCG TTA GTG GCT GAG TAT CTG AGC ATG CTG GAG GAC CAC AGC TAC
leu val his ala ala ala leu val ala glu tyr leu ser met leu glu asp his ser tyr
4801/1601 4831/1611
CTG CCC GTG GGC AGT GTC AGC TTC CAG AAT ATT TCT TCC AAT GTG CTG GAG GAG TCT GTG
leu pro val gly ser val ser phe gln asn ile ser ser asn val leu glu glu ser val
4861/1621 4891/1631
GTC TCT GAG GAC ACC CTG TCA CCT GAC GAG GAT GGG GTG TGC GCA GGC CAG TAC TTC ACC
val ser glu asp thr leu ser pro asp glu asp gly val cys ala gly gln tyr phe thr
4921/1641 4951/1651
GAG AGT GGC CTG GTA GGC CTC CTG GAG CAG GCC GCG GAG CTC TTC AGC ACG GGA GGC TTA
glu ser gly leu val gly leu leu glu gln ala ala glu leu phe ser thr gly gly leu
4981/1661 5011/1671
TAT GAG ACA GTT AAT GAG GTC TAC AAG CTG GTC ATC CCC ATC CTA GAA GCG CAT CGA GAA
tyr glu thr val asn glu val tyr lys leu val ile pro ile leu glu ala his arg glu
5041/1681 5071/1691
TTC CGG AAG CTG ACA CTC ACT CAC AGC AAG CTG CAG AGA GCC TTC GAC AGC ATC GTT AAC
phe arg lys leu thr leu thr his ser lys leu gln arg ala phe asp ser ile val asn
5101/1701 5131/1711
AAG GAT CAT AAG AGA ATG TTT GGA ACC TAC TTC CGA GTT GGT TTC TTT GGA TCC AAA TTT
lys asp his lys arg met phe gly thr tyr phe arg val gly phe phe gly ser lys phe
5161/1721 5191/1731
GGG GAT TTG GAT GAA CAG GAG TTT GTC TAC AAA GAG CCT GCA ATT ACC AAG CTT CCT GAG
gly asp leu asp glu gln glu phe val tyr lys glu pro ala ile thr lys leu pro glu
5221/1741 5251/1751
ATC TCA CAT AGA CTA SAG GCA TTT TAT GGT CAA TGT TTT GGT GCA GAA TTT GTG GAA GTG
ile ser his arg leu glu ala phe tyr gly gln cys phe gly ala glu phe val glu val
5281/1761 5311/1771
ATT AAA GAC TCC ACT CCT GTG GAC AAA ACC AAG TTG GAT CCT AAC AAG GCC TAC ATA CAG
ile lys asp ser thr pro val asp lys thr lys leu asp pro asn lys ala tyr ile gln
5341/1781 5371/1791
ATC ACT TTT GTG GAG CCT TAC TTT GAT GAG TAT GAG ATG AAA SAG AGG GTC ACA TAC TTT
ile thr phe val glu pro tyr phe asp glu tyr gln met lys asp arg val thr tyr phe
5401/1801 5431/1811
GAG AAG AAT TTC AAC CTC CGG AGG TTC ATG TAC ACC ACC CCG TTC ACC CTG GAG GGG CGG
glu lys asn phe asn leu arg arg phe met tyr thr thr pro phe thr leu glu gly arg
5461/1821 5491/1831
CCT CGG GGA GAG CTG CAT SAG CAG TAC AGA AGG AAC ACA GTC CTG ACC ACT ATG CAC GCC
pro arg ala ala leu his thr phe tyr arg arg asp thr val leu thr thr met his ala

A

5611/1871

5581/1861 ATT GAA GTT GCC ATT GAA GAC ATG AAG AAG ACC CTG CAG TTA GCA GTT GCC ATT AAC
ile glu val ala ile glu asp met lys lys thr leu gln leu ala val ala ile asn
5641/1881 5671/1891
CAG GAG CCG CCT GAT GCA AAG ATG CTT CAG ATG GTG CTG CAA GGC TCT GTG GGA GCT ACT
gin glu pro pro asp ala lys met leu gln met val leu gln gly ser val gly ala thr
5701/1901 5731/1911
GTA AAT CAG GGA CCA CTG GAA GTA GCC CAA GTG TTT TTG GCT GAA ATT CCT GCT GAT CCA
val asn gln gly pro leu val ala gln val phe leu ala glu ile pro ala asp pro
5761/1921 5791/1931
AAA CTC TAT CGA CAT CAC AAC AAG TTG AGG TTA TGC TTT AAG GAA TTC ATC ATG AGA TGT
lys leu tyr arg his his asn lys leu arg leu cys phe lys glu phe ile met arg cys
5821/1941 5851/1951
GGT GAA GCT GTA GAG AAA AAC AAG CGT CTC ATC ACG GCA GAC CAG AGG GAA TAT CAG CAG
gly glu ala val glu lys asn lys arg leu ile thr ala asp gln arg glu tyr gln gln
5881/1961 5911/1971
GAA CTC AAA AAG AAC TAT AAC AAG CTA AAA GAG AAC CTC AGG CCA ATG ATC GAG CGG AAA
glu leu lys lys asn tyr asn lys leu lys glu asn leu arg pro met ile glu arg lys
5941/1981 5971/1991
ATT CCA GAA CTG TAC AAG CCA ATA TTC AGA GTT GAG AGT CAA AAG AGG GAC TCC TTC CAC
ile pro glu leu tyr lys pro ile phe arg val glu ser gln lys arg asp ser phe his
6001/2001 6031/2011
AGA TCT AGT TTC AGG AAA TGT GAA ACC CAG TTG TCA CAG GGC AGC TAA
arg ser ser phe arg lys cys glu thr gln leu ser gln gly ser OCH

GAAAAAGCCATCTTCATTCTGGAGACTGTGGCCCTGCAACCCCTGGAGAAGGACTTGTGTTACTTAAAAAAATGGGACATT
TGCCACCCAGGACTGACTGTACACTCCCTGATCAGCCAGCACTCTGGAGCTTGGGATCCCAGGAACCATGGAATTATT
CCCAAATGGACTCTGACCAGATTTTGCACATACTGGGGGTGGCGGGATGGAGGATGGGTACTCAGGCATGACTGCGTAT
TTATTAAGTGTGTTTCCACAATGTACCAAACAAGGCATAAGCAGCTCTCTGCTGACTGGCAATCACTGCCCATC
TGAGAGATGATTCCTCTGGCCCATATTGAATTATTGGAGTAACTCAAATTGCTGAGGAAAATGGAAAATTATCC
ACCAGTCGATTCAAACCTGAATTCACTCTTATAGGAAGGCAGGGCAAACCTGTAGGAGTACGAAACATTTCATAAAAT
CTACAAAGGGAAAGCCTTACTACAATTCCAAAATCATCATGGTTGGAATTGGGAGGAGATTATTGTGAACCTGTTAC
CTTTGGTAATGGGACTAATTGCTGTATAGTTTTGTTTATTACTGTACATTAATTAAACATGCATTAA
TAGAAGAATACTCAAAGCACTGATGTAGGAGATACACGGTACTTGGAGCAGTCAGCCAAAATCACAGATACTGCTTT
CACTTAAATGGAAACAAATTCTCGATAATGCTTGTCTTCTTATGTCACTCTGTGACTATCTATTTCCTCCTC
TCTGGGACCAAGTTCTTTTATAAGCAATAATCTGTGTTTCAATTAGAACATTTGCTGTCAGCATATGT
ATATCAGCTACAAATATATTCAACTTTGACTCTTGACAAAGGACTTTAGGAAAAGGAGGAACAAAGACATTATTG
AGAATTAAATTATATTTAATATGACTGTGACCTGACTGATAATAAAGATGTAATAAGAATTGCAAGCTAAAAAA
AAAAAAAAAAACTCG

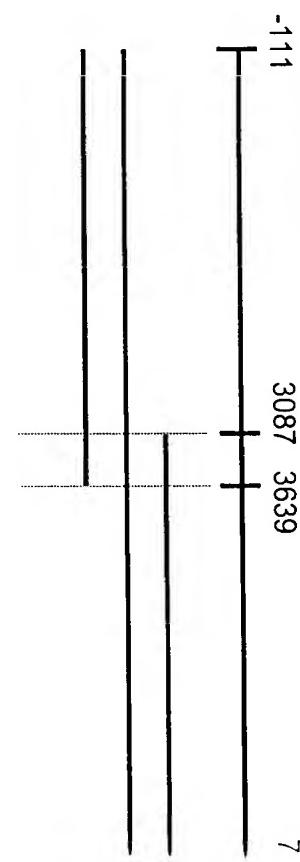
A. Allelic variations: single nucleotide changes (polymorphism) between CLASP-5 cDNA isoforms

Isoform	Nucleotide(s)	Consequence
1	1727	C to T change; mis-sense mutation changing codon from alanine to valine
2	1749	A to G change; silent mutation
3	2277	G to C change; silent mutation
4	2853	C to T change; silent mutation
5	3427	A to G change; mis-sense mutation changing codon from lysine to glutamic acid
6	3777	C to T change; silent mutation

B. Alternative splices

Isoform	Difference	Nucleotide(s)	Consequence
1	exon deletion	1806-1944	premature, in-frame stop codon leading to the production of a truncated, most likely soluble protein
2	exon insertion	between 2857 and 2858	additional, in-frame 48 nucleotide exon that contains a stop codon at the second codon, which would lead to a truncated, most likely soluble protein

nucleotide numbering
as in Figure 6A



552 nt overlap

1st partial exon (nucleotides 3793 to 3952)

CCAGCTGTCAGCCAAGCTCAGTAACCTTCAACGCTCATTCCATGAGGCTAG
AGTTCCTGAGAATCCTCTGTAGCCATGAGCATTACCTCAATCTGAACCTTTT
TTTATGAATGCTGATACTGCTCCAACATCTCCTGTCCTCCATATCTTCCCAG
GTAATAAAAGAATTATTAACtAAAAGAATTATTCAAGCTAT

2nd exon (nucleotides 5809 to 5948)

GCTCATAAAATGGCTCCTTACGTTCTGTAGAACTCAAGCTCCTGCTCCAGCT
TCCAGGACCAGAAGATGCCAGCATGTTGATCTGACTTCCGAGTACCGCCA
GCAGCACTTCCTCACCGGGCTCTTACAGAACTGGCTGCTGCCCTGGATG
CCGAAGGGGAAGGGTATGTTCTGGCATTAAATGGAAGATGAAGC

3rd exon (nucleotides 13662 to 13831)

CATAACCTCTTGATTCCCTGTGTTGTGCCAACAGAATCAGCAAAGTACAAAGG
AAAGGCTGTCAAGTGCATTACAGCCTGTAAGTTCTCACGACCTGGACCCAC
GCTGTGTC~~AAACCAGAGGTGAAGGTCAAAATCGCCGCCCTTACCTACCTT~~A
GTTGGCATCATTGGATGCTTGCACAGCTGTGACTTACAGGTAATGG
CCCTCTGTTTCTTCTGGATTG

4th exon (nucleotides 16948 to 17087)

TGTTGACTTGACATCACAAACGATGTTTCATTGCAGT~~TG~~CAGATACTCGCA
GATACCGCACCAGTGGCTCGGATGAAGAACAAAGAAGGAGCCGGTGCATT
ACCAGAATGTGGCTCTGCCATAGCAGGGATAATTCAATTGAAAACAAG
TGGAATAGTGTGCTTCCTGGTATGTTGGTGCACATGTGTCTGGTTATT
TCAT

5th exon (nucleotides 19281 to 19463)

TGGCCTCCATCCCCAATCTGCCCTCAGCCCTATAAGCAGTACAACATG
CTGAAACGCGGACACTACTCGCAACCTCATGATCTGCTTCTGGATCATGAA
AAATGCTGATCAGAGCCTCAITAGGAAGTGGATTGCTGACCTGCCATCAACG
CAGCTCAACAGGATTAGATCTACTTTCATCTGTGTTATGTTGAGTAT
AAGGTAAGTCTGGAGTGGCACAACTTATACCAGC

6th exon (nucleotides 19829 to 19958)

CACCAAAGGACATGTCCTCCTACCTCTGTCTTGTCCAGGGAAAACAGAGTTCT
GACAAAGTCAGTACCCAAAGTCTGCAGAAGTCAAGGGATGTCAAGGCCGG
CTGGAAGAGGGTTGCTGCGTGGGAAGGGGCCAGAGGGGAGATGATGCC

7th exon (nucleotides 20928 to 21015)

TCAAATTCCCTATCATGCATTTCTTAACCTCTAGGGAACGACCGATTCCAGGC
CTAAATGAAAATTGAGATGGAAGAAAGAGCAGACACATTGGCGGCAAGCT
AATGAGAAGCTAGATAAGTGAGTCACTCGGCAACTTCTGCTACTTTACCT

8th exon (nucleotides 25765 to 25861)

GCTTAATTGACCTCTGTTCTCTAGAACAAAGGCCGAGTTAGATCAAG
AAGCCTTGATCAGTGGCAATCTGGCTACAGAAGCACATTAAATCATCCTGGA
TATGCAGGAAAACATTATCCAGGTGAGGAAAACAAACACCCAATCTGATTG

9th exon (nucleotides 27242 to 27376)

GGATTCAATGATGCTGTTCTTCCATTCCCCCAGGGCAGCTCGGCTCTGGACTG
TAAAGACAGCCTGCTGGAGGTGTTCTGAGGGTGCTGGTGAATTCTCTGAAC
TGTGATCAGAGTACCAACCTACCTGACTCACTGCTTIGCAACACTCCGTGCTCT
CATGCCAAGGTAAACTGGGATGCTGTTCTCCTTTAATT

10th exon (nucleotides 28582 to 28734)

AGTGATGCCTAATGGCCCTTATGTCCTCTAGTTGGAGACTTACTCTTCG
AAGAGGAGGTGGAACAGTGTTCGACCTATGTCACCAAGTCCTGCACCACTG
CAGCAGCAGCATGGATGTCACCCGGAGCCAAGCCTGTGCCACCCTTACCTC
CTCATGAGGTTCACTTGGAGCCACCAAGTGTAAAGAGTTCAAACCAAGCTGAG
TGACCTGGAATCAG

11th exon (nucleotides 31046 to 31204)

TTACTTCATCTTTTTTTTTCACTGATGCAGAATTGCAAGAGTAAAGA
TGCAAGTAACCATGTCCTGGCATCTTGGTGGAAAGAGCACCAAGACTTAA
TGAAGAGCACCTGAGAAGATCCTTGAGGACAATTGGCCTATTAGAAGAG
GACACAGCCATGCAAGATGACTCCTTCCCACCCAGGTACACCGAAGCACAT
ACCTTGTCTCATGCATGAGT

12th exon (nucleotides 32755 to 32855)

AGCTAAGATTATTTGAGGCTACACTTTGCAGGTGGAGGAACCTCTCTGT
AATCTGAATAGCATCTTATATGACACAGTGAAAATGAGGGAAATTCAAGGAAG
ATCCTGAGATGCTTATGGATCTCATGTACAGGTAAAGCTTCCCTGACACACTCA
AGGGACACCATT

13th exon (nucleotides 33663 to 33855)

TCCTCAAAACTACTTCTCACTCAATCTGTCCTCAGAACATTGCAAGAGTACCA
GGCATCTCCTGATCTCGGGCTGACCTGGCTCCAGAACATGGCAGAGAAACAC
ACCAAGAAGAAGTGTACACGGAGGCTGCCATGTGCCTGGTGCACGCCGCTG
CGTTAGTGGCTGAGTATCTGAGCATGCTGGAGGACCACAGCTACCTGCCCGT
GGGCAGTGTCACTCCAGGTAGGGTGTGTGCAGCTTCCCTAGAGCAGTG
GTTC

CTGTTCTCCAGGCTTAACTGTGGTCTCTTCAGAATATTCTTCCAATGTGCT
GGAGGAGTCTGTGGTCTCTGAGGAACCCGTACCTGACGAGGAATGGGGTG
TGCGCAGGCCAGTACTCACCGAGAGTGGCCTGGTAGGCCTCCTGGAGCAGG
CCCGGGAGCTTCAGCACGGTCAGTGCCCAGAGGGCATCCGGGCCTGGC
C

15th exon (nucleotides 40166 to 40297)

AATTCTCTGTATGCTCTCTCTTCCAAGGGAGGCTTATATGAGACAGT
TAATGAGGTCTACAAGCTGGTCATCCCCATCCTAGAACGCGATCGAGAACATC
CGGAAGCTGACACTCACTCACAGCAAGCTGCAGAGAGCCTCGACAGCATCG
TTAACAAAGGTAGCCGGGAGCCTGGCTGGCAGGTCTGTTAC

16th exon (nucleotides 40755 to 40889)

TAAGGAGAGCTTTTATAATTGTTCCTCAGGATCATAAGAGAACATGTTGGAA
CCTACTTCCGAGTTGGTTCTTGGATCAAATTGGGATTGGATGAACAG
GAGTTGTCTACAAAGAGCCTGCAATTACCAAGCTCCTGAGATCTCACATAG
ACTAGAGGTAAAGAAAAGTGAATTCTGTGCGCCTGACCTGGTACACTTAC

17th exon (nucleotides 42307 to 42396)

AACCTTATAAAACTGTTGGTTCTCTTACCTAGGCATTTATGGTCAATGTTT
GGTGCAGAAATTGTGAAAGTGAATTAAAGACTCCACTCCTGAGAACAAAACCA
AGTTGGATCCTAACAAAGGTATAACAAAAATTACAAAAACTAACCATCAAGC

18th exon (nucleotides 45250 to 45486)

TCTTCTCCCTCCGTGCCTTTCCCCCTAGGCCTACATACAGATCACTTTGTG
GAGCCCTACTTTGATGAGTATGAGATGAAAGACAGGGTCACATACTTGAGA
AGAATTCAACCTCCGGAGGTTCATGTACACCACCCGTTCACCTGGAGGG
GCGGCCTCGGGGAGAGCTGCATGAGCACTACAGAAGGAACACAGTCCTGAC
CACTATGCACGCCCTCCCTACATCAAGACCCAGGATCAGCGTCATCCAGAAG
GAGGAGGTAATGCACCCAAAGGGATTGCCACCACTGGATGAGT

19th exon (nucleotides 48664 to 48807)

ACAGTGACTCCCTATGTTACGTCTCATGTTAGTTGTTGACACCGATTG
AAGTTGCCATTGAAGACATGAAGAAGAAGACCCCTGCAGTTAGCAGTTGCCAT
TAACCAAGGAGCCGCCTGATGCAAAGATGCTTCAGATGGTGTGCAAGGCTCT
GTGGGAGCTACTGTAATCAGGTAAAGCAAAACCAAGAGGTGGCAGCTCCT

20th exon (nucleotides 50892 to 50998)

TATATTCTTTTTTTTTTTTTCCCCACCAAGGGACCACTGGAAGTAGC
CCAAGTGTGTTGGCTGAAATCCTGCTGATCCAAAACCTATCGACATCACAA
ACAAGTTGAGGTATGCTTAAGGAATTGAGGTAAAGAAGGAAAATG
GCTGGGAATTTCAGTAGAG

21st exon (nucleotides 62398 to 62568)

TCAAAAAGAACTATAACAAGCTAAAAGAGAACCTCAGGCCAATGATCGAGC
GGAAAAATTCCAGAACACTGTACAAGCCAATATTAGAGTTGAGAGTCAAAAGAG
GTAAGAACAGGGCAGAGGAGGCCTTCCTGTGGGAT

22nd exon (nucleotides 63040 to 63294)

CCTCCCTCTTTCTTAATTCAGAGGACTCCTTCACAGATCTAGTTCAGGA
AATGTGAAACCCAGTTGTCACAGGGCAGCTAAGAAAAGCCATCTTCATTCGT
GGAGACTGTGGCCCTGCAACCCCTGGAGAAGGACTTGCTGGTACTAAAAAAAT
GGGACATTGCCACCCAGGACTGACTGTACACTCCCTGATCAGCCAGCACTC
TGGAGCTTGGGATCCCAGGAACCATGGAATTATTCCCCAAATGGACTCTGA
CCAGATTTTGCCATACTGGGGGTGGGGGATGGAGGATGGGTACTCAGGC
ATGACTGCGTATTTAAGTGTTTTCCACAATGTACCAAAACAAGGCCAT
AAGCAGCTTCTCCTGCTGACTGGCAATCACTGCCATCTGAGAGATGATTTC
CTCTGGCCCATTTGAATTATTGGAGTAACTCAAATTGCCTGAGGAAAA
GGAAAAATTATCCACCAGGCTGATTCAAACTGAATTCACTTACTCTTATAGGAAG
GCAGGGCAAACTGTAGGAGTACGAAACATTTCATAAATCTACAAAGGGA
AGCCTTACTACAATTCCAAAATCATATGGTTGGAAATTGGGAGGAGATT
ATTGTGAACTGTACCCTTGGTAATGGTGGACTATTGCTGATAGTTAT
TTTGTTTTATTACTGTTTACTTAAATTAACATGCATTTAAGAAATAC
ATTCAAAGCACTGTATGTAGGAATACCGGTACTTTGGAGGCAGTCAGCCAA
ATTCACAGATACTGCTTTCACTTAAAATGGAAACCAATTCTCCGATAATGCTTTGC
TTTTTCTATGTCACTTTGTACTATCTATTTTCTCTCTCTGGGACAA
GTTTCTTTTATAAAGCAATAAATCTGTTTTCATTCAGAACATTGTGCTG
TCTGTCAGCATATGTATTCAGCTAAAAATATTCAACTTTGACTCTCTTTG
ACAAAGGACTTTAGGAAAGGAGGACAAAGACATTATTTGAGGAATAAATTT
ATATATTTTTAATATGACTGTGACCTGTACTGATAAAAGATGTAATAAGAA
TTGCAAGCTAAAAAAAAAAAAAAAAAAAA

1 GTTCTCTGGTTAGTCACTTAGTGAATTAGATAAGTTTTCCAATTATGGGCTTAATTCTCAGTTAAATAAGAAGGGGGGG
92 TTGAGACATTGAGCCGTATCAACGAAAAGGATAGGACCATAAAAGCAGTGACATACAAGCTCATTGAGCAGCACTGGACAGGGTTA
183 CATAAGAGCGAACCCCTCCAGCATGAGAACAGCCATAGGCCCTGCAGTGAGGAGGGGACATCCAGAGGACAGGGGAACTCCAGGGG
274 AGAGGAGGATAGGGCAGAAGCTTATAGATCTGGGTAGGCTGCCACAGCACAGTGAGGAGCTCTGGGTAGAGAGCTCCAAGGGCTG
365 TAGCAGCTTAGGCCCTGTATCTGAGGCTATCTTATCTAGGAAACAGCTGTCATCGAGAATTATGGGTCTGCAAAGGAGGCA
456 GGCACAAATGGATGAAAATCTGCTTATAGGCTATTITGTTAATGACTGGCATGGTAAAATTGAGTTAGGTATGAGCCAAGGGATC
547 AGCACCAGTCGACAGTGAGAAAATAAACAACCGAGGGCAATACACAGAAAACATCTTGACTTGTTACATAACACCTCTCCAAAA
638 AAAGAAAAGACATGACTTTCTAAAGGAAGTCCTACAGCCATACCACCTGAAGCAGCAGCTCATCTAAGGGAAAGTCCTGATA
729 AGGAAGCAGAATAAATAAAAGAATCTAAGGAAACATCAGAAGTCCCAGCATCCCCATCCCATGACCCCTGCCCCGAGCG
820 GATCTCTGCCCAGGACCCACAGAGATAAGATGGCAGAGGGACACTTCACTCTCCCTTGCCCTCCCTTCAGTATTAGGATTCAGGT
911 TAGTCTGTTTCAAACATTAAAGTGAAGTGTATGAAAGATAACCTAGATCACCACCTGCAGACCCAGGCTAGCTACTCACATGAGGCTAT
1002 CTTCATCCGACAGGAAATPAGGAACCCAAAGGTAGATATCTGCCCTCCCTGCAAACACTCATGTTTATTTCTCTTCTCTTCTT
1093 TCTTTCTCTGCTACTATCTTCTCTTAAATCTTCCCTTTCTTGTAGTACCATTTCTTTTGTCTCTCCTGACTCAAATG
1184 CCAAAGATCCTAGACCAAGGGAGTCAGAACCCCTGCATGCTGATTCTGGGACACCTTAAATTACTCTGATTTGGGAGCATGCCCTTG
1275 AGGCATTCATTCCTCATCTGCAAAAAAAACTAGGCTGGATTAGATTATCCACTGATTCTGCTGCTGCTGCCAGTGACATCCACCGA
1366 TGTTACTTAGCACCATTAGTGGCACTCAGGCCCTGAAGGTCACTGACCCCCATTGCTGGTAGTTAAATTCTGATCCAGCTCTAGA
1457 TACAGGATGAACTTCATACCTTAAGCAAGGTGATTCTTACAAAGTCGCTGACTTATCATTTGCTACAAACCTTATGTTCTGCT
1548 CTAAAGATTGGGAAATGGATATAACCTTAGTCTTTTAAATAGTCACTAAAATAGACAAATCATTCTCAATTACTGCTGTCAC
1639 TTACAATGTGCCACTAACACCCATCCCTGGTAGGGCATTGCGGCTTGACCTTGGGAGGAGTTTGGGATGACCTTA
1730 CTCCCTTTCAAAACATCACTCAAATTGTTTGAATGTCATTTCCTTTGAGATGGAGCTTACTCTGTTGCCAGGCCAGAG
1821 TGCAGTGGCACAATTGGCTCACTGCAAACCTCCGCTCACAGGTCAAGCGATTCTCTCCACCCCTCCGGAGTAGCTGGGATTACAG
1912 GTGTGTGGCACCCTCCCGTAATTGGCAATTGGTAGAGACAGGGTTTACCATGTTGGCAGACTGGCTTGACTCTGCCCTCA
2003 AGTGATCTGCCCTGCCCTCCGAATGGCTGGATTACAGGTGTAGGCCACACCCAGCTGAAATGTCCTTGTGATGATGCT
2094 CAAGGGACCATGAAAGATACCCATAGTGGGCTCTTTAAAGTGCCTGAAAGTCTTCATGTCCTTGTGAACTGGCTTGACCCGAC
2185 ACCTGTTAATGAGTAACCTAAGTGACAGGCACATGACCAAGTCTTCATGTCCTTGTGAACTGGCTTGACCCAAAGTCAGCCAG
2276 GAGTGACACATCGAAAGGTTATGGATCTGTAACGTGCTTACATAGAAGTCATGTTGGTTAAATATATAATGCAATTAC
2367 TTATTAAAGTCGATGCTAACTATGAAATTCTGAGGCAATTGTCCTCACACATGGCTTCTGAAATGGCTGAAATGAG
2458 GTTCATTAAATATATTGTTAGTCGTAATTCCAGCACTTGGGAGGGCAAGGGGGAGGATCACCTGAGGCCAGGAGTTGAGACAG
2549 CTGGGCAATATGGTAAACCCCATCTCTACTAAAATACAAAATATTAGCTGGCATGTTGATGTCACCTGTAATCCAGCTACTCGG
2640 AGGCTGAGGCAAGGAAATCGCTGACCTGGGAGATGAGGCTGAGTGCAGGAGATCATGCCACTGCACTCCAGTCACAGCAAG
2731 CTCTGATCT
2822 TGTGTATATATATATGTTATATATATATGTTATATATATATGTTATATATATATATATATATGTTAAACAAACTACTCT
2913 TGAAGTGTAGAGAGGGAGGAGCATGTCATCACAGGATGTCCTCATATTCTGCTGGGTGGGGGTGATAATGATGAAATTGAG
3004 GAGCTCAAGGTGCATCAGGCCCT
3095 CTGGGGAGGAGGGCT
3186 CTAACACTCTCTAGCTTCT
3277 GTACTTGCTAAACAGTTGATGCCCTCTCTAGGCACTGGCTACAACTTCCATGGGCTTAAATTAGAAAATACGAGGAGAACAC
3368 TTGTTAGTAAAGTCACTAGTCATATTCTATTAGTTTACACAGGCTCATATGAGGTCACCTCTTCTTACATGTTCTGAAACATAGT
3459 AGACTTTGTTAAATTAATATGTTACGGTGAAGGATGACCTCAAGATTTCAGCACAGGCTCCATGGTATAAAGATTGATAAAGT
3550 GGAATCGGATAAGAATCTCATCTGGACAGCTACTAGGCTAGAATCACTAAGGGGAAAGCTAATGAATGAAATTATTAGACCTCTC
3641 TGTAATGCGAAATTACATCTGTCACATCCATTAGTTTATTAAATTGATGCTTAACTGGAGAGAAGGATATTGTTGTC
3732 TGCCACCTCAACTCCACTACCTCTGTAATAAAATGTTCTTCT
3823 AACCTCATTCATGAGGCTAGAGTCTCTGAGAATCCTCTGAGCCATGAGCATTACCTCAACTGTAACCTTTTATGATGCTGAT
3914 ACTGCTCAACATCTCTGTCCTCCATCTCCAGGTAATAAAAGAATTATTTACTAAAAGAATTATTCAGCTTACAGCTATTCA
4005 TAGCTAGTTAATCATGTTCTATAAAGGTTAGTCATTTGACAAAAAAATCAAACAACTCAACAGGATCAAGTATGCTACC
4096 CTGAAGTTACACCACTAGCTAAGAATTAAACATCTAAGTAATTGTTCTCCCAGGCTCAAGGCTCCCTGATCAGGTTAAGTAAAGCCAA

4551 TCA GTAGCT CAG AGC AT CAG GCA AAA ATT TCT CATT ATT AGG TT AT ATT CTG TT GCA T ATT CCTT GAT ACT TAG TAC AAA AGT GAA GG CT TG
 4642 TCT TA CT A ATT GAAA AAA ATT CCT AGC AT AT A TG CCA T AT GGC AT GAT CC AGA ATT AGC T AC AT GAC CA T CT ACT GT GAA CAG GG AA
 4733 AGA T CT GACT CACA ACC AGC AA TT CAA AAT GT TAA AACT TTT GCT CC AT CCT GCC CAG GCT TAA GCT CTA AAA ACT CCT GGA CTA AG
 4824 CTT ACC CT AGG TT CTT CCC ACCT TCC CTT CT GACT GCCC CAG GAG TGG CCA AAC CAG GGG CCA CAG GCT CAC AAA ACC AT GA AGG AT TT
 4915 CTAA AGA CAC TTT AGAT GCT CTT AAT GAA AT ATA AAG TGT GCT CC CAG GAT ACA AAT ACAG GAC AGG AA TT ACT GAG GAC CGG TAA AT CTA
 5006 AT ACT TCC CTC CT GAC AT CACT TGT AGT TCC AGGG CAG CAA AGT CT GCA AT GT GCT TAA GCAA ATT CAG AAG TGT GAG CT GAG GCG G
 5097 GCAC GGT GGC TCA CATT GT AAT CCAG C AT CCT CGG AGG CCA AA AGG GAG TGG AAT ACT TGT GAG GAG GAT TAC CAG GCT GAC CA AC AT GA
 5188 TGAAA ACT CAT AT CTA CTT AAAAA AT GCA T CAG GT GT GGT AGT GT GACT GT AAT CCT CAG GCT ACT TGG GAG GCT GAG GCA T GAG GAA
 5279 TTG CTT GAA CCC CGG AGG GAG GAG GT TGC AGT GAG CAC AT GCA AT CAG C CT GGG TA ACAG GAG TGA CACT CTG TCTT CAA
 5370 AAAAA AAA AAA AAA AAA AGA AGG TAC CT GAG TA AGC AGGG CTT AA ACA AAG GGG CA TTT GGT ACAG GGA ATT GCG TGT CTC CAG C
 5461 CCC AGG ACT GGG TGG GGT CTT ATT CTG TCC CAA CCT TT AT GCT AGG ATT TT AT CCT CAG CTT TGT AAT TCA GGT GGG AA AGG GAG G
 5552 TCAT CATT TCA ACAG ACC CCG GG GT CTT ATT CT CATT CCT CGT GGG CAG TCT CTC TCT GT GGG TAT GGG AA AGT GCA GAG TAA CC
 5643 ACT CT GT GGC CT CAG TTT CTT GT CCG CTT AT CT GGC ACC AGA T ACCT AC ACT CCT CACT GGG GAG GCT CAG GAG CT GG TGG GAG G
 5734 CTC AGT AAA ACT GGG C AT CGC T ATT TT CATT CCAG TCT TT GT GGT CTA AAA AT GGT CTT ACAG TTT CTG TGA AACT CA AGC TCC TGC
 5825 CCAG CTT CAG GAC CAG AAG AT GGC CAG C AT GT T GAT CT GACT CCT CAG GAG CAG CACT CCT CAC CGG CT CCT CTT CAC AGA
 5916 ACT GGG CT GCT CCC TGG AT GCG CAG GGG AGG TAT GGT TCT GGC ATT TAA AT GGA AGA T GAG CCA AAA AAG CAG AT GT TCT TTA A
 6007 AAAT TT GCA GT CAG TCT CAC ACT TGG TAAA AACT CTA CT GTAG TGT GACC AGT CT GAG GAG TGA AAG AC AT CT GT CTT GAG AAT AT GG
 6098 TAC CCATA AGC ACA AGG CACA AGA AAG GCT TT CTT GT GTAG AAG AGG CAC CAG GAT GGG TA AGA AACT AC AAA AT GACT TTT CT GGT CAA
 6189 CT ATT CAG TGG AATT ACC AGT CT GCT AT AGC AGG TT CCA AGG AT GCT TTT GAT TGT GAA CTC CCT CAG GAG CAA AGC ATT TT A AC
 6280 AAAGGG AT ACC AT GCA GAG GCA ACC CAA AGA TGT CACT TGG TCA AAG CT GAT GAG GAA ATA AT GGC TGT GAG AAG GCA CCT GT C
 6371 TGCC CAG AT TAG GT TCT TGT CAG AC AGT GCT TCT CAG GCA AAG GAC ACC AC AT CCT CAG GGT MCT GTA AT CCT TTT ACCT CTT AT C
 6462 TGAT TACT CAG GGA C AT GT GGC GAG TAC TAC GCT TA GACT CA AAG CAC AGA GAG GAG TGA GAC CAC CCCC TT CTT CTC TCC TAT GTC
 6553 CGT GAG CT CAT GG AGT CAG AAA ACC CAC AGC C TAT CT GAT TGG AGT GAA AAA AGA TAT GCT CTA AAA AT TAT TCA TCC GTT CA A
 6644 TATT GAT GCT CTT CCA TGG GAC AGA CAC AGT TGG GT TGA GAG TAC AT CAG AGA TAT AT CAG TGC AC AAA AAG CAG ACAA ATT GCT G
 6735 TCA AGG CTC ACC AT CCA CAG GAG CTC AGT CAA AGA TGT AGG AT TGC CAA AGT GAG AT TCA AC ACT CCT GT GGT GTT AGAT TAGG C
 6826 TAC AGA ATT GCA GAG GACT TG CCA CTA IT CCT CAA CAG GAG CAA AT CTA AAG GCA TGG CTT CCA GAG CTT CTC CAG GAC CCA CTC
 6917 CT AT ACC AA ACC AT CCT CTC GCA CAG GGA ACCT GT TCC CAT GACT CCT AT GCT TA GCT AAG AGG TT CCA GAG AAT AT GT TGT CTA ATT AT GGT
 7008 GC AT TGT GT TAT GGT AT GGG ATT TCT AAT GGA ATT CCT CAA GTT CT CAG TCT CAA AGT GCT TCA ATT TGT GAG GAA AC AGG AGG
 7099 TT CCT CCA TGA A A CCA TAC AGG GAC AT CG A CAG T C G T G C A C T G C A A C A G T C A C A T G T G T C C A A G T C A T G T A C C
 7190 TGA AGT GACT TAT GT CTC TCT GGG CAGG AT AGT GT GGG GAA TCT AT GCT TTT AGT CT CTA ATT CTG CTC CTC CAG A G T A A C T G
 7281 GCCT GAG TATA CCT TGT GAG AACT TCA CCT TGT GGT CTA ATT CT AT CCT CTC CTA CACT TA ACT TAA ATT CCT CCA ATT AT CT C
 7372 GTAG AAT GTT CT ACC CTA CACT A A C A T A A T T C C A C A G C A A C A A A A G T G C C A C G A A C A C T T A G C T A A A G T A A G T A C
 7463 AAG TT GAA C A C A A G A A A A T T G C C T C A T G C A C A A C A T G T A C A T A C A T T T C T A G T G T C T T T A A T G C A G T A C T T A A G T G A I T T C
 7554 TAG A A C A T T T G A T G T A T T A C A A T A G C G T A G T T C T A T T T C T A T T C T A G T G T C A A T G C A C A T T C T G T C T G G A T T C
 7645 TTG TAA ATT GT GGT GGT TAT ATT ACT GCA CAA AGG T A G G T A T A T C A G A G A T A A T C A T G T A A T T G T T A A T C A G G T A G A T A C A I T C
 7736 TT CCA A C A C A C A C A C A C A C A C A C A C G C A C A C A C G C G C G T G C A C A T G C A C A C A C A C A G T A C C C T T C T C C C C C A A A A G G A T A I T T A
 7827 A C A G T G T A A C T C C T T G G C A A C A G T A A T C C C T G A T T G T G G G G T G C A T T A C T C T C T G G A A G T C A T T A G A T A C A C T G T C A C A A C T T
 7918 C A C T A T G T A A G C T A A A C C T C A G A A A A A T G T C A T G C T C A C A A T A C T T C T A T T G T C A G C A C T T C C T T T T T T C T T G C T T C C A
 8009 A G A C A G T T G G C A G A C T T G C C A A C A C T G T A C A T C C A T G G G C A A G G G C C A G A A G G G C G C A T C T C A G T A C C T G A A C C C T A G G G A G C T A C A G
 8100 C T C C A G T T C C A C T T G G G T T C C C T T G G A T C T G C A C A C C T G G G T A T A T C C A C A A G T G T C T A T G C A T C G T T A A T C C A C A G T G T T T T C A A G
 8191 C A T T T G A A A A T T G T C T A T C A T C T T T G T G T A A G C C A G C A A T T G C A T A G C C T A A T T C A A A C C A C A C C A T T G C C G G A C C A C A G A C T A A G
 8282 A A C A A G A A A A C T T T G A A A T G C A A T T C A A A T T A T C T A T T G C C A C A G T G C A A G G T C G A G T C A T T A A A A T T T G G T T A A T A T T
 8373 T C T A T A A C G T T G A A T C T G A T G T G A C C T T A T T T G T G T G A A A G A A A A A A T G T A T A T A T G T A T A T A T T T T A A A T C T G G A T T C T T
 8464 T T T T A T T A T A T C A T A C T T T A A G T T C T A G G G T A C A T G T A T A C A A C A T G C A G G T T G T G T A C A T G T A T A T A T G C C
 8555 C A C C C A T T A A C T C G T C A T T T A C A T T A G G T A T A T C T C C T A A T G C T G T C C C T C C C A C T C C C C C C A C C C C A T G C A C G T C C C C G G T G T G A T G
 8646 T T C C C A C C T A T G A G T G A G A C A T G C G G T G T T G G T T C T G C C T T G C G A T A G T T G C T G A G A A T G C A G G T T T C A G C T T C A T C C A T G T C C
 8737 C T C A C A C A C A C A T G C A C T C A T C C T T T T T A T G C C T C A T A T G C T G T C T A T T C T G C T G T C T A T T C T G C T G T C T A T T C T G C T G T C T A T T C T G C T G T C T

9101 AGATAAAAAATATAATATTCCAITATCCTCTCTCAGTTACAGGCCCCCTCAAGTAACCACTATTCTGACTCTTATTAGAAATTAATA
9192 CTGCCCTGTGTTGAATTTCATAGTCTTTGCTGAATGTTGACTGTGACATCCTTAGTGTTGCATGTCAGTATTCAITTTTT
9283 CAACTGTATGTCATATCCTTGTCTACTATAATTCTCTCTGTAATTGACATTTGGGCTGCTTCTATTGTTGGGTTATTGGTATTAT
9374 GAAAACAGCTGCCGTGAACATGCCCTGCTGGATTTGGGGACCTTGAAGACTCATTTCTTGGGGCTATAAATACAGCCTTTTTAT
9465 TTAAATATACTGCTCTGAATAGTTAATAAATATGTCATGGCTTAAACAAATGTCAAAAGAAATATACTGAGCTAGGAAAAGAAG
9556 AGAAAACAGTCAAAGCAGGAAGATGCCAGGAATAACAAAGGTGATAGCCAAAATAATGAAATAAGAATAGAAAACAATCACGGAAA
9647 TCCCGAAGATGAAAAGCTTGTATTGAAAAGAGCAACAAAATTCCAATCTTAGCTGAGCTGACCAAGAAAAAGGAAGAGACTCAA
9738 TTACTAAAATCATATTGAAAGATTCAACACAATCATATCACAAGAGACCTTACAGAAATAAAAGGATTATAAGAATACGATGAACAA
9829 TTGAAAGCCATCAAATTGATAACCTGATTAATGATAAACTTCTTAAAGGTACAAGTACTAAATTGACTCCAAGAAGATATAGAAA
9920 ATCCAATAGACCTACAGAAGTAAAAGATTGAGTTAGTAATCAAATTCCCACATACACCTACTATGTAACCCACACAAATTAAAATTAA
10011 GGCTGGGCGAGTGGCTCATACCTGTAATCCAGCACTTGGTGGCAAGGCAGTGGATCACCTGAGGTCAAGAGTCAAGACCAGCCT
10102 GGCCAAGGTGATGAAAACCCGCTCTACTAAAATACAAAATTAGCTGGTGTGGGGGGAACTGTAATCCCATCTACTCGGGAGGC
10193 TGAGGCAGAAGAATCATTTGAGACTGGAAGGCAGAGGTTGCACTGACCCAAAGATCATGCCAATGCACTCCAGCCTGGCAACAAGAGCAA
10284 ACTCCATCACAATAATAATAATAATAATTATTTAAAATTAAACTTCTTACAATAAAAGCTCAAACCTGGGGGCTTACTGAT
10375 GAATTCTACCAAATATTAAAAGAATTAACTCTAATTTCACCAACTTCCAGTCTCTCTTCAACGAATGGAAGAGGTGGAATACTT
10466 CCCCCACTTGTCTATGAGGCTAGCATTACCTTATACTAAACAGACAAGACATCATGAGAAAATACAGGCCAGTATCTGATGAATATAG
10557 ATGTAAGACCCCTAACAAACACTAGCAAACAGTGAATTCAACAGCATATAAAAGGATTATACACCATGGCTAAGTAGGATTATCTCAGGAA
10648 TGCAACATACCCCTCCMCTCCTGAAATTCAATTGTTGACCATATTAAATAAAAGGACAAAACCCATACAATCTTGTAGATGCAA
10739 AGAAAAGCAATTAAATAAAATCTAATAACGCTCTGATAAAAACACTCAACAAACCTTTAGGAATAAGAGAACTTCCCAACTGACTT
10830 AAGGGCCTATGAAAATCCACAGCTAATGTCACATTAGTGAAGGTTCTATAGGAGTTCTATAGGCAATTAGGAAAATAAAACAAAGACATCTAGGCGAGCGTGC
10921 CTACCCCTGCCACTTCTATTCAACATATTAGGAGTTCTATAGGCAATTAGGAAAATAAAACAAAGACATCTAGGCGAGCGTGC
11012 GTGGCTCACGCTGTATCCAGCACTTGGGAGGCCAAAGTGGCACAGATCGTTGAGGCCAGGAGACTGAGAACACCTGAGCAACATG
11103 GCAAAACGCCATCTCTACAAGAAATACAAAATTAGCTGGCATTTGGCTTGTGTTGAGTCCAGCTACTTGGAGGTTGAGGCTG
11194 AGAATTGTTGATCCAGAAAGCGGAGGTTGAGTGCAGTGCAGATCACCTACTGCACTCCAGCCTGGGCACAGAGTAAGACCCCTGTCT
11285 AAAAAAAAAAAAAAGAAAAGAGAAAAGAGAAAAGGATTAAATTGAAAAGAGTAAACTATCTTATTCATAGGTTGCATAAGTCAACAA
11376 TCTTGTTTATAGAAAACCATAACGAAATCCACAAAAAACTCCATTACAACATAATAATGAAATTGAGCTGAGTGTGATGGTATAAGTCAACAA
11467 TACAAGAATCAATTGTTCTATACACTTACGATGAGCAACTGAAAATGAAATTAAAGAAAACATTCTATATAAAATGACATCACAAAG
11558 AAAAAAATTTAGGAATAATGTAACAAAAGAACACAAGAGTTATACACTAAAATGACAAAACACTGTTGAAAGAAAGATATAAATAAA
11649 TGGAGGATATCATATGTCATGAATCAGAAGACTTATTATAAAATGCAATACTCCCAATTGATCCATAGATTAAATGAGTTCTTCT
11740 CAGAACTCTAGCTTGTCTTTTTTTGGCAGAAACTGCAACTGCTTAAACTGTTGAGGAAATTCAAGGGACCCAGTATAGCC
11831 AAAACACCTTGGAAAACAAGAACAAAATTCGAGACTCACACTTCCCAATTTCACAAACTTACTAACAGAAAAGTAGTCAAGACTATGG
11922 GTTCTGACATATGATAGACATATAGATCAATGAAATTGGGTTAAGAGTCAAAATAAATCTTATATTAGTCAATTGTTGAGGTTCTTCT
12013 AAGAGTGCACAAACATTCAATGGGGAAAATAGAATTTCATTAATGGTGTGGGACAATGGGTATCCACACTCAAAGAATGAGTT
12104 GGACCCCTATATTACACTGTAACAAAATCAACTAACATGAAATAGATCAAGACCTAAATGTAAGAGCTAAAACATAAAATTGTTACATAAAAT
12195 TATAGAGGTAAATCATCATAGACTTAGAAAAGGAGCTGGTTCTTAGATATGACACACTCGAAAGTAGTGTGAGTACAAGAAAATAGATAA
12286 CTGGACTTCTGAGTAAATTAAACTTTGTGATTATAGGACCATCACAAAAAATGAAAAGGCAACACACAAAATGGGAAAATTGGC
12377 AAATCAAAACCTAATAGGGACTTGTATGGTATATATTAAAATCTTCAACTCTGTAATAAAAGACAAAATACTCAGTTTTT
12468 AAAAGGCCAAAGATCAGAATAGACATTCTCCAAAGAAGATACAGCCTAAAGACCATGAGATGTTGAGCATCATTACCGTCAGGGAGAT
12559 GCATATAAAATCACATTAAACCACTTGTGATACCCACAGAAGATGGATATAATAAAAAGACAGGTAAATAAGTGTGGCAAGAATAAAA
12650 TGGAGTCTCAGACACTGCTGGGGATGTAATTGTCAGGCCACCGTGTGAAAACAACCTGCTGATTCTCTAAAGTTAACAGAGGCC
12741 TGGGGCTCGGGCTCACGCCCTATATCTCAGCACTTGGGAGGCTGAGGTGGCAGATCATTTGAGGCCAGGAGTTGAGACCCCTG
12832 GCCAAGATGGTGAACCCCTGCTCTACTAAAATACAAAATTAGCCAGGTTGCACTGAGCCAGGAGATCTGTCATTGCACTCCAGCCTGAACAACCTCATGCAA
12923 GAGTCAGAAGAATTGTTGAACCCAGGAGGTTGCACTGAGCCAGGAGATCTGTCATTGCACTCCAGCCTGAACAACCTCATGCAA
13014 AAAAAAAAAAGTTAAACAGACAGTTACCATACAAAGCCAGCAATGTAACAGGTTATGACCCAAAGAAAAGTAAACATTAAACTTGTG
13105 ATACACATACTCATGAGCTGGTAAAGTCACAAATAGCTCAAAGCAGAAACATCCAAATGTTATGAGTGAATGGATAAAATC
13196 ACCAATGGAATTATTAGCAATAAAAGGAATGAGTACTGATGCTACAAATGATAAAACTTAAACATCATGCTAACAGCAGACCC

13651 TGTGCCAACAGAAATCAGCAAAGTACAAGGAAAGCTGTCAGTGCATTACAGCCGTCAAGTTCTCACGACCTGGACCCACGGCTGTGTCAG
13742 AACCGAGGTGAAGTCAAAATCGCCGCCCTTACCTTACTGTTAGTGGCATCATTGGATGCTTGCACAGCTCTGTGACTTACAGG
13833 TAATGGCCCTCTGTTTCTTCTGGATTGTTGGGCCCTGCCAAATGCCCATCGRAATGAGATCTGTCACTCGTCCAGTGC
13924 ATATGTCATATGAGCAAATCTTCAACAAAATGGTCTCAAACCTTTAACACTGGCCCAATCAGTTAAAAAACAACAAAAGTCACAT
14015 AGCCCCATATGTTTTGTTAGAAAATTATTAACATTGTTCTCAATTCAATTCAATTGATCTGTAGTATGGAATGAAC
14106 TTATAAACCATACACACAAAGGATGAAAATAACAAGAATTACTGCCTTGGAGATCGTGTCAACACCTGAAATAAGGTGTCCT
14197 TCCCCTTCTTATGAAATGTAAGCAAAAGGGTCTCAAATGACCCAACTAAATGCTCATGCAAAAAAAATCATGGCATATGT
14288 CAATATTAGAAAATTCTATCATAGCAGAGAAATTGGCCCTGGCACTAAGTGACTAATTGAGGTTAATATGTCAGATTGACAAAG
14379 ATGAAAGGATAAGTGTCTTAAAGGGTTTGTGAAACCCAGAACTATTTACAAATTACATTGTTAGGAATGTAGAGGGAAAA
14470 AATCTCCCAAATAGAGAACTTCTATGTTGACCTGAGAAAGGCACTTCAATTGGCTACCCCTCCACAGCGTACGTGTCACGTG
14561 TGTGCGCGTGCACACACACACACAATTCCATCCCAGAGTAGTCTGTTTCTCCAGCTCAAGGGAGTTGGAAATG
14652 TCCATGCTCTCTGGCAGCTGAAGGGTCTGGGATCTGGCAATGCTCCCTGAGCAGCTCTTATAGTCAGTCATCCCTGAAATC
14743 TTGGAACCATCAGGTTTCCACCTCTCCAATGGAAAGTGTCACTGTGACTAAATGTTATAAATACATTACGGTACAGTATT
14834 AGTTCATGAAAATGGACCAAAGCCTATAGTCTTAACTACAAATCATAAAAGAACAGTATTCTGTGAAAGTAACTCTC
14925 CTTTCTGAAAAAAATCTCTAGCAAACGCTAAATGGAGTAACCTCAGAGAGAAAGGGTAGAGCACGGGTTAGGGTTGAGAGTCTGTT
15016 GGGGCCAGAGTGCAGGAAATGATAAAACAGCTCAGCACACAGGGTAGAGCAAGCTGACCAGGTTAGGTACAGCTCTGAGCG
15107 CCTGATTGCGGGAAITGAAAATGGAGTTTAAACATTACAGACATAATGAGGACATGGCATGTGACTTGTAGCCAAATTGAGAA
15198 TCCTTGTGAAAGTTCTAAAAGTCCATTCAAGGCTTGTGTTCCAAAGCCAGGAACACTTACGACACATTATCGGAAGT
15289 TCTTGTCAATTAGAATATTCCGTATCAGTGACTCGAACTCAGAACATTTCACATTGGCTCCAGCCTTTAAACCTCAAAGACTCTT
15380 TGTATATGCGGTGAACTTTATGATCATGATGTTGATGAGGAAAAACAGTGAAGGTTAGGTATAGGTTACACAGTACAT
15471 CTATTAAATATGAACTTTAGGTGATAAACTCATAAGAATACAGTCTTCAAAAGATGCACTGAAATTCCAGGCCCATTAT
15562 TAATATAAGTGACCTTGAAGGCTCAACCTTTCCGTGCTTCTCTGTAAACTATAAAATGTGATAACGATGTCACCTTTAG
15653 GTTACGATAAAGGTAAAGACTAGAGAACTTATAAAAGCATTGGCAGCACATGTTGATCTTCCCGATTTCCTGTAATTAA
15744 CCCTATGGCAATTAAAGGGAAATAGAATGTCCTATGTTAGTGTGATAATGTTACAGGTTGCTGCAATTTCAGTGTGTTTAT
15835 TTAACCATTTGTTAAATCCCAATGGCTTAATAAGTATTCTAAATGGGTTGGCTCGGTGTTGGCTCACCCCTGTAACTAAA
15926 GCACTTGGGAGGCAAGGTGACAGATCATTGAGGCCAGGGGTCAAGACCACCGCTGGCAACATGGCAAACCCCTGTCTCACTAAA
16017 ATACAAAAGTTAGCCAGCGTGTGCCACCTGTAAUTCCAGCTACTCGGAGGCTGAGGTAGGAGGATGGGGCAGAGATGCACTG
16108 ACCGGAGATGGCCCACTGCACTCCAGCTGGATAGTGTGAGACTTGTCTAAAATAAAATAAAATAAAATAAGAACGGTT
16199 TTGTTTCAACACTTTAGTAGAGAAATGTTATTACACTCAAATTCTAAATAATTGAAAGGCCAGATGGCTGAAATGTCACAGGTC
16290 TAGAAAACATGAATTAGGAAACATGTTGAAATCACTGAGGAACTTGTGAAATGTTTACACACTTTCCAGAAGTCTGAAATA
16381 TAGTTTGTGTTACTGTTGAGCTTAAACAGACTTTAAACGGTTATGTTTCTGTTTCTGCAACATCTAGACCTATCTA
16472 TAAATATGGGTATGTCATATTACATAGATAAAATGACATTCAATTAAACATTGAAATATGACAATTATGCTGAATTGTC
16563 TAATAGCTTATTCTATATAAGGTGTTGCTTAATTCTGTAATTGTTAGGTGTTCTATTGTCATTAAATAATGCAATTAT
16654 TAGAATGAGATGAAGTTAAAATAATTGTTACAGGAACAAATTGAAAGGTTGAAACATTTTGTCACACAATTGAAAATTAA
16745 TTCTAGCTTAATTCTGTGCTAGACATTGCTTGTGTTTCTTCTGTTAGGTGATAGCAGATACTAATGCTAAACATCAGGTT
16836 TGAAAATACTGTGCTGACTTTAGTGACTGAGAAGTATCAGTCTTATTGGTAGGGACATGGGAAATGTCATGTTGACTTGACATCA
16927 CAAACGATGTTTCAITGCACTGAGATACTCGAGACATCGCACCCAGTGGCTGGATGAGACAAGAAGGCCGGTGCCTAAACCA
17018 GAATGTTGCTCTGGCATACGGGAATAATTCAATTGAAAACAAGTGGAAATGTCGTTCTCTGGTATGTTGTCACATGTGTC
17109 TGGTTGATTTTCAATTGCTCTGTCTTGTGTTTGTGTTGTTGAGACAGAGTCTACTCTGTGACAGGCT
17200 GGAGTCAGTGGCAATCTCAGTTCACTGCAACCTCCGCTCCGGTTCTCATGCTCAGCTCTGATTAGCTGGGATTACAGGGT
17291 GCACCAACACCTGCTTAATTGGTATCTTACTGAGAGACGGGTTTCCCATGTTGGCAGGGCTGGCTCAAACCTCCAAACCTC
17382 AGATGTTACACCCACCTCAGCCTCCAAAGTGTGGGATTACAGGCTGAGCCACTGTCGCCAGCCTCTCTGTCATCTCTGATGGGAGA
17473 ACTCAGCTTACAGCAGGCAATTGATTATTCCGCTCATTTGCTGGAAAAGAATCAATTGAGGCTGGCAGCTGAGGTGACACAAGTCAGA
17564 AAGGGTTCTGCAACAGGAAGAGATGTGCAAGTCAAAATTCAAGGCACTGCCCAGTGGCTGAGTTCAAGTCAAGAAACCGGAAAGAGGAA
17655 AACTGGGAAATCTGAAATCTGAAAGGAGATCAAACCAAGGCTCTGAGGGTGGAGGCTTGGTAGCTGTTAGCTGATGTTGTC
17746 ATTCAACAGTGTAACTGCTCCAAAAAAAGACTGAGGAGGCTGGTAGCTTAAACAAAGACTGCAAGCTTCACTGTCAGTATGTC

22751 TTTGCCCTGGCTGTAAAAGTGGTGTGAAACCTTCCCTTGGCTGCAGGAGCAACACAATATTGGCCTTACCCACCGCATGGCCT
 22842 ACCGGMTGCTTAGGTTAGTCTCAGCCCTGCGTACCTGGTGGCTGCTGGTTGTCAGAGCTAATGATTCAACATCTACAAATTCTCAGG
 22933 GGCTTTGCTTTTTTTTTAAAGGAAACCATAGAAAAATTATAACTGGTCAACCTACCTGGAAAATTATAATTATAAT
 23024 AATGGTGCTTCCACCTCTTCTGATCAATTTCGATTATTGGCTCTGTTCTTCTAGAGCAATGATCTCAACTTACGGCTGCACATT
 23115 AGAACATCACCTGGGGAGCTTAAAAGCTGTCAGTGCCTAGCAGCTGCACCCCTAGACTACTTCAAATGGAAACTCCAGCAGGACCCAGAAATCAGT
 23206 ATTGGTAAAAGTCCACGGTATTGCAACCAAATTGCAATTGTTAAAGGAATAGATGGTAAGCTTTCTTTCTTTCT
 23297 TTCTTTTTGAGACAAGGCTCACTCTGTCACCCAGGCTGGAATACAGTGGCACAATCACAGCTCACTGCAGCCTTACCCCTCGGCCT
 23388 CAAACAATCCTCCACCTCAGCCTCCAAGTAGATGGGACTACAGGTGTGACCAACACTGCTTACTTTGTAATTGAGAGAC
 23479 AGGGTTTACCATGTTGCCAGGCTGGCTAAAACCTGGCTCAAGTGATGCCACCTGTTAACCTCCCAAAGTGGGGATTACAGGC
 23570 ATGAGCCACTGCCCTGGTCAAGATGGTAAAGCTTTAAAACAGATTAGTGTAGTGTAGGGTACAATATGAATGACTAACACTA
 23661 CTGAACTGTATACTACAAATGGTAAAGATGGTAAATTGTTAGGTGACTTTACAAATACAAAAATTGGGAAAAACAGAITTA
 23752 GGACACTCTAGAMPTGCTGCTCAATACAGTAGCCATTAGCCACATGTGACTATCAAATGCTGAAATATGGCTAGTCAATTGAGATA
 23843 ATAAATCAAGATAAGTGAAGAAATACGCCACAGATTGAGGCTTATGTGAAGAATAAAAGAATAAAATATTCACTAGTAATTGTTATATTG
 23934 CTTACATGGAGACAGTATTGTTATTTAGGATAACAAATATAATTAAAATAATTCACTTTCTTACTTTTAATGGTTAC
 24025 CACAATATAAAAATGACATATGTGGCCCCATTGTTCTACTAGACAGCATTGCTCAAATTAAAACAAACTAGCAGTCATAATTAA
 24116 AAAAGGATGATAAGGGCCGGCACAGTGGTCACACCTGTAATCCCGACACTTGGGAGGCCAGGTAAAGGGATCAGGAGTCAGGAGA
 24207 TCGAGACCATCTGGCTAACCGGGAAACCCGCTCTACTAAAAATACAAAAGTAGCCGGCGCGGGTGGCCGGCGCTGTAGTCCC
 24298 GCTACTCGGGAGGCTGAAGCAGGAGATGGCTGAACCGGGAAAGTGGAGCTGGGTGAGCCAAGATGCCACTGCACACTCGC
 24389 AGCCCTGGGCAGAGGGAGACTCGGCTCAAAAAAAAAAAAAAGGATGATGGGTTAAATGGTAAATTGATGTTAGTGTAC
 24480 TTTATCAGATATAAAATTGATGGCTCATGCCCTGGTCCAGATCTCAGGAGGCTAAGGCGAGCATCTGGCCAGGAGITCG
 24571 AGGCTTCAGTGAGCTATGATGTCACCTGCACTCTACCGCTGGGTGATAGGCAAGACCCCTGTCTCAAGGAAAAAAACTT
 24662 TTAGATTCAATTATTACACATATACTACTGGAAAATGAGAAAAGTGTCAAGTGGCTTGGGACAGAGCCCTATCCTAAACAT
 24753 GAAAACAAGTAAAACACACAGAAGTACTTATTGTTAGTCTCAGTGTAGTAAAGCCTGCAAGTGGCCATTATTAGGTTAATGGG
 24844 CGCAAGAACAGGTAAGTGGTAACCCCTGGCCAGGACATATGAGCTGATAGGAAACCCATGGTAACATCTGGCTACTGAG
 24935 CATCTGGTAAAGCTCAATTCTTCATCCCTTCTGCAACTAGATTGGTGTAGTAAACAAATATCCCTTACAGCTTCACCTAGA
 25026 TTTCATAAGAATGGTGGCTAGCAAAACATTCTGATTCTTAAAGTGAACCTAATAGGAGGAGGAGCACCCAGGAATTCCCAT
 25117 TCCCCTAGTGGCAATGACCACTGTCGGAGGATTAGTACACCTGCCCTACAGGAATACAGTCTGTTCCCAATGGAAAAGCAG
 25208 AAAGACCCACGCACTGGTGAGCAACCTCAAGGTGATTTGGGAAGTTAAGGGCTACTCTCCCTGACTTGGCTCTGAAGCTCCAC
 25299 CCTTTCTCAACCCAGCGCATACTGCCAAGTTCACTCACATCCAGTCTCACTCTCTCGCATTTGCAAGGGCTTCCCTACTCAGA
 25390 TCTAGAACATGGTTCTGCGTAAATACGCCATGGCCCTGAGCTTGGGCAACATCTGAGAGGGAGCTTAATGCTAGTAGCAACT
 25481 ATAATCTCTCTACCCATAGTATTGTTATAATTGTCCTCATCAATTATAATAGTGGGATGAGGATGACCAGGAACCTTAC
 25572 CTAGACAGTTTCCACAAGACATGAATCACAGAGGACCTGACTGTAGTTACTCAGGGCAGTGTGCTCTGTTCTAAGGGTGC
 25663 ACTATTGGAGATTCTTACCTGGTGTAGATGCCATGAGGAACTGGCCAGGTTGACATGGCTTTAATGACCTCTTG
 25754 TTGTTCTCTAGAACAAAGGCCAGTTAGATCAAGAGCCTGATCAGTGGCAACTGGCTACAGAACACATTAAATCATCTGGATATGC
 25845 AGGAAAACATTATCCAGGTGAGGAAAACAAACACCCAACTGTGATTGTTGGCATGAATATGTTACTAGAATAAGGACTCTTATGC
 25936 AATTGTGAAAGACATAATGTGATCCCATAGTACCTTTTAAAGGAAAGTGAAGTGTACTATTACACAGTGTACCTTATA
 26027 AATTCCAGAGATACCAACATTCTCTGGCTCTTGTGACTTAGGCTTACTTGGAGAGGGTAGGTGTTGGCAGCTGACCCCTTGGTT
 26118 AAATCTGTTGAGTATGTACCAAGTTATAATGGATGTTGGTTATGTTAGTATGACAGTACTGGTAAGTGAAGTATTCT
 26209 GATGGGTCAACTCCAGTGTGAAATGATGGTCACTGTCGATATGGGAGTATGATTAGTACTAGGCTAGTAAAGAGGTGAAATTGACA
 26300 AATGATGTCCTCTGGACTTAAATTGTTAAGGAAAGTCATTGACCATGAAATTGCCATCCCTGCTGAGAAAATATGGTTGTGA
 26391 ACTTTGTCACATCAATTATCTCTATGTGATTATTCCACAAATTCCAAAGCTGCTAGGTTAATGAGTTTAAATACCCCTGAA
 26482 AAATGAGTTCTTACATGTTCCATTGAGAAGTCATTCACTAGAGTAGGTCAGGATTGCTTTAGGGCTAGAAGAAATATCGTGAACAC
 26573 AGTGAATCTTAATCTCAACTTGTGAAATTGCTAAAATCAAGTAATCATCATAACAAAAAAACACAAAAAGTATGTGATTTTGT
 26664 TGACTTAAATCTTGTGATAACTTAAATGCTGGTACATTTACCTTACCTTATGTTATAGCACAATTAGGTGCAAATATCTACTA
 26755 GCCCCAAATATAATTGCAAGTTCAAGAAGAGCTGAAACCTTGTGTTATGCTTGTAGGCTAGAAGAAATATCGTGAACAC
 26846 CTTCTATTAGAAACATTGCTGCCACCAAGCAGCCCTGTTGACTGGGAAGGCCACAATTGTTGCTTGTGATCCCATAGGAAAGCTATGT

27301 GAATTCTCTGAACGTGATCAGACTTACCCACTGACTCACTGTTGCAACACTCCGTGCTCATGCCAAGGTAAACTGGGATGC
27392 TTGTTTCTTCCTTAATTAAAGAGTAAGATTCTCATCTAGCTCATCTCTCTTCAGGTGACCAAAGTCACAGAGCATATTAAAGT
27483 GGCATCACAGTAAAGGCTTAAGTCTTCTTAGGAAGAAAGCAGATGCCCTGATTCTGTGGGAAGGCCATGGAGAGGAAAAGCAGTGGCT
27574 CCCATATTGAAGTGTCGACCTAACCTCTAGAAGTTAAATGGCATTGCTGAAGGTCTATGACATGAGAACAGAGATCACTGAGTGAC
27665 TTAGCAATTCACTCTTCTCTGATAACCTCTGACTGAGATAAATCTCTATGTGACGCCATTAGCTTACAAAAAGTCATGCCA
27756 TAAAATGCCAGGAAGGTAGAGAAATGAATTCTCACGGCTGAGGAATGAGGATTACCTGGGTAAACATGAGGATTATTTTCTTATT
27847 TATTTTATTATTTTGTGAGACTGAGTCTCGCTTATGCCAGGCTGGAGTCCAGTGTACCATCTAGCTCACTGCAGCTCTGCC
27938 CCTGGGCTCAACCGATTCATGCCCTAGCCTCTGAGTATTGGGATTATAGGGGTGTGCCACCCACCCAGCTAAATTGTATTAG
28029 TAGAGACAGGGTTTCAACCATGTTGCCAGGCTGGCTTGAACCTGACCTCAAGTGTACCCACCTGCCCTAGCCTCCAAAGCAGTGGGAT
28120 TACAGCGGTGAGCCCCCGTCTGGCTTATTTCCCTTATTGAAGATCTCAATTGGTGCCTTACATGGGTCTTTAAATTAAAAAA
28211 GTAAAAATTCTCTGCTCATCCTCTCAGGACCATTTCTTCTCATCACAGTAATTCCAGAACCCAAGAAACTCAGGTTCTT
28302 CCATCATAGTTGTGATTTCACAGTGAATGCCACCTGGCTCACAGTGCAGTTGATAACACAGCTGACCCCTTATGGCAGACGTTTCAATT
28393 ATTAATCTCAAGTCACTCCATTGTTAAATCCATCTCTGAAITCACATAGCTATTATCTTATGGAATAATGCAATTAACTCTCTAGG
28484 CTTTTGCTTGCCAAATGGACATTGATATTCAACGGTCCAGAAAGTGTATAAACTGCCAAGTGTACGCCCTTATGTCT
28575 CTCTCTAGTTGGAGACTTACTCTTGAAAGAGGAGTGGAACAGTGTCTGCACCTAGTCACCAAGTCTGCACCAGCAGCAGCATG
28666 GATGTCACCCGGAGCCAAACCTGTCGCCACCCCTTACCTCTCATGAGGTTGAGTTGGAGCCACAGTGTAAAGAGTTCAAAACAGCTGAG
28757 TGACCTGGAATCAGTAGAGAAAAATTGATGTAAGACATCAGCTGCCAAAAAAATAAGGAATTGGCAGTTAGCAGTTACTCTGTCC
28848 TGTGAGAAAGAACAAATTGAGTATGAGATAGACAGCTTCTATTGATCTAAAAGTGAATTCTACAGATAAAATGAGTGT
28939 CTCTATCAGTGTGTTCTAAAATAGACGCCAGGGCCAGGAACGGATGGCTTACCTATAATCCAGCACTTGGGAGGCCAGGTGAGT
29030 GGATCATTGAAGTCAGGAGTTCAAGACTAGCTGGCCACGATGGAAATCTGCTCTACTATAAAATACAAAATAGCCAGATGTGTG
29121 GCGCATGCTGTAATCCAAGCTACTGGGAGGTGAGGAGAGAATTGCTTGAACCTGGGAGGCAGATAATGCACTGAGGCCGAGATTGCC
29212 CCATTTCATCCAGCCTGGACAACAGAGTGAAGACTCATTTCAAAAATAATAATAAAATAACCAGGTGAGTGGCTCAIGTC
29303 TGTAATCTAGCACTTGGAGGCCAAAGCAGGACATCAGATGGCCAGGACTCAAGACTAGCCTGGCAACATGTTGAACACCCCGT
29394 CTCTACATACAAAAAAATTAGCCGGGTGTTGTCACACACGCCGTAATCTACTCTGGGAGGCTGAGGCACGAGAACTCCTGAAAC
29485 CCAGGGAGCAGGGTGTAGTGAACCAAGATTGCTGCCTACTGATTCAGCCTGACCCCTGTCTAAAAAAAGAAGAAAGAAT
29576 GGAAGAGTATTAGATAAAAGTATCATCTGTTGGGAAAAAAATAACATAAGACAGGTTAGAATTGAGAGTGTCTGTTCTGTTCTAAA
29667 TTCTGACTAGCTAGTGCACGAAATGACCTGTCAGGAGGATTITAATGATCGGTCTCATCTAACCTGAGTTTATTAAATTATTAA
29758 TTTATTATTAGAGACAGTGTCTGCTCTGTCACCCAGGCTGGAGTTAGTGGCACTATCAGAGCTACTGCAGCCTTCACCTGGCTC
29849 AAATGATCCTCTACTTCAGCCTCCATGTAGTAATCTGAGGATTACAGGCACGAGCCACCTTACCCAGTAATTTTTGCAATTGTTG
29940 AGACAGGGTCTGCTGATGCCAGGCTGGCTAGAACACCTGAGCTCAAGTGTACCTCCCTCAGCCCCCAAAGTACTGAGGATTATA
30031 GGCATGCCCTCTGCCATGCCAACACTTGAGTTTATTCAAAAGCTACGAAGACTTGGAGTTCAGCTTATTATGAAACAGTCAGTT
30122 GCTTGTAGTTGCTAGATTGATACCTCTTGGAAATTCTCATTGTCAGGACTATGAGCTTACAGGCTCAGCTTGGGACCTTGAGGTACCTGTT
30213 TGGCCCATGGAAAAAGTCAGCCTCTCCTAAATGATTAGGGATGTTATTAAAGACATTCTCAGGGGACCTTGAGGTACCTGTT
30304 TTCCATGGGCTGTAAGAAAGAAACAAAACCTTGTGCTTACCCGGAGTCAAAATCTCAGAAATGCTGCCACAGAAGTCCCCTGA
30395 TTTATTATTAGAGACAGGGTCTGCTCTGTCACCGTGGCTGGAGTGGAGATCTGGCCACTGCAACCTCTGCCACCGT
30486 TTAAGCAATTCTCTGCTCAGCCTCTGAGTGGAAATTACAGGTGTCCACCCAGCTGCAATTCTGTTATTAGG
30577 CGGGGTTTGTGTTGCCCCAGGCTGATCTGAACTCTGAGCTCAAGTGTACCCCTCTGGGCTCCAAAGTGTCTGGGAGGCTGAG
30668 CCAGAAGTCCCCTTCTTAAATAAAGTTAAATAAAGTCCCAAGAGAAACTCTGGCACAAAGGATAACTGTATTCTGGACCCAACT
30759 TTATAAGAATCTTCACTGGCTGAGCACAAAGGCAGGCCAGTCTCAATGAAAATTAAAGGGAGCTGACAGATTATGTGAGGCAATGT
30850 CCATTAAACCAATTAAACAACAATATGAATGTTGCAAAGTGTAGCTCCATTCTATTGAGAGAAGAGGAAATAATTAGACGGGCAA
30941 AGGAAACACTGAGGAGTTGTTGTCCTGGCATGCTGCTTCACTTATCTACTGCTAAGTGTGCTTATTAACTCTCATCTTTTTTT
31032 TTTCAGTGTGAGAAATTGCAAGAGTAAAGATGCAAGTAAACCATGTCCTGGCATCTGGGAGAGGACACAGACTTAAATGAAGA
31123 GCACCTGAGAAGATCTTGGACAGGACAATTGGCCTATTCAAGAGGGACACAGCAGTCAGATGACTCTTCCACCCAGGTACCCGA
31214 AGCACATACCTTGTCTCATGATGAGTTGGGATGCTCAACTATTGTTGATGATGACATATACACAAATTATATAAAATAT
31305 ATAACATTGCAAGTATTATGTCATGAGTGTGCTCAGCTGAGGTTTAAAAGAAATACATCCATCACTGCCCCAG
31396 TCTCAATGCTAACAGTCAATTGGAGTGTCTCAGTCTCAAAACAATTAGGAGGCAAGTACAAGACAAGTGTACATAAGTGC
AAACTGTGT

31851 TATGGTAGGGCTTAATTAAAGTGAGAATGAATGAATGAATGTTTCTTTATTCCCTTAGTCAGAACCCTATTCCGGAGTCGTGTTTG
31942 CCCTCTCGGGCAATGTGCATCTCATCTCAGAACCTTTCTCTCTACATCAAATCGTCTCTCATTCATCTGACCACAC
32033 TCCTTCTAGTTCTCATTTATAGAGGTATTACCAACAAGCCAATCAAACATGGGAGTTAATAAAGGTCTCAGTGCCTCACCC
32124 AATGAAATGACTCTAGTGGTAAAATTAGGAGCCCTGGCAAGCTGGCAGGGGAACGGGATAAGACAACATTCTGGCTGAGTTAC
32215 CTGCCAGGGCTCTAGATCAAGCCATAGTCTCTCCCTGGTACTGCAGGCTCCCTGGACCTCCACTGTTGGTTATAAAGGAA
32306 TAAATGATTACAAGAGGTTCTAAATCTGAAGCCCTGGGAAGATCCAGGGCTCTGAGACATGGAACGACTCAAGTGAAGTCTAAGCT
32397 GCTTCTACTTGGTATAAAAATCCCTGATATTCCAGAGTAGAGTTAGAACTTTCAGGTACAAATAACTGAAACTGGTCAAACAA
32488 TAAACAAAATGTTGGAGATAGAGATAGTATGGGATTAGGCAAGTTCAGGCTCAAGGGCAACACACATCCAGGTCTCATAAACTC
32579 CTGGGCTGGCAAGATGAAAATCACAGTCAAGGCTAGATGACTCCGCATCCCTAGTCTCTGCTGCCCTGATCCAACCCCTCCATG
32670 GCTTCCCCTTGCATTCAATGAACTTCACTTCACTGGACTACAAGCTAGATTATTGGGTTACACTTTGAGGCTTACACTTTGAGG
32761 GAACCTCTGTAATCTGAATAGCATCTTATATGACACAGTGAATAATGAGGAATTTCAGGAAGATCCTGAGATGTTATGGATCTCATG
32852 ACAGGTAAAGCTTCTGACACACTCAAGGGACACCATTGGGGTCTGAGGATTTCAGTGTGGATTCTACTAATGTAATGATCAGC
32943 TAACATGGATATAGTGTATTGGATGATGCAAAATAATATAGAAATTAAACATTCAAGTGTAGGTTAATCATATGTAAGTTTCAGAAG
33034 GATCTCTTAATTAAAGTGGGCAATAATATGTTATTAAATAACATTATAAAATAACACTATCTTCACTTACACAGAA
33125 TCAGAGAAAAGATAATTAAACTGTTAGGAAATTATTTCCATAATGCTTGTCAATTATATTCACTACATACATATAGTTACTTIA
33216 GAAATCATTTTACCGCTAGTTAAAAGTGGTCTTCAGTCCTGTTCAAGGAAATTACTACAGAGGCAAATTGTTGATGCAAC
33307 ATCATAACAAAGGCCAGTACTTTGCTCTGTTTATTGAGAGAAAGGAAAGGAGGAGAAATTGCTGAGGCCATTAAATAGA
33398 CATCATGTTACAGGTATTTCCTCCATAAGGCTTTACTAAGTACTATTCTGGAGGTACGGACAGCTTAAACATGGAATTAAATAG
33489 TTGGACTAATAATGTTCTCTGCTGTTCTGGAAATATAGGGCAAATCTCAGGGTGGAGGGTACAGGGAACCTTGGGAGAAA
33580 AAAAGAAAAGGTACACAAAGTAGAGAAGAACAGTGTCAATTACCCACTGCTCTCAAACACTTCTCACTCAATCTGCTTCAGAATTGCA
33671 AGAGTTACCAAGGCATCTCTGATCTGCGCTGACCTGGCTCAGAACATGGCAGAGAACACACCAAGAAGAAGTGTACACGGAGGCTGC
33762 CATGTGCTGGTGCACGCCGCTGCGTTAGTGGTGTAGTACATGAGCATGTCGGAGGACACAGCTACCTGCCGTGGCAGTGTACGGCTTC
33853 CAGGTAGGGTGTGCACTTCTTAGAGCAGTGGTCTCACTGGGCATTGGTCCCCAGGCCAGGGACATTGGCAATGTCT
33944 AGATACATTGGTATCACAACTGGGATGGGTGAGTAGGTGCTACTGGCATCTGACTGGTAGAGGCCAGGATGGTGAACACATTCTGC
34035 AATAGGAGAGCTCCCTGACAAAGAATTGCTGCCCCAAATGTCGTAGTGTAAAGGTTGAAAAATTCAAGTCTACATCAATTGTC
34126 TTCTCTAATTGCTTCCCATCGCTGGTTTTTTAAATTACTGTTACAATAATGCCACTAGCCATTAAATAGCATTGACTTTAGGA
34217 GACATTGCAAACACTTCACATGCTCATTTGACCTCCCGTAAAGCTGTAGGCCAGGTAGGTAGGGAAAGGGGTATTATTC
34308 CACTTCGGGATGAGAGAACTGAGAGAGCAAGTTCTAAGGTCACTTAAACTCTTTCAAGGACTGTAGTTGACACAGTACTGACA
34399 TTGTGAAAATTGGAAAATGGATAATGATTCTCTGGCCCTTCAATTGTTGATGTTCACTTCAATTGTTGACCTTCAAGGTTAGGAAAGG
34490 AATCCAAAATCAAGGAAAGAATCAAATTGACCTGAGAAGGGAGACAAAGGCTGAATCAGTACCTTCAAGGTTAGGAAAGCTGAAATTACCA
34581 CAATTGGCAAAGGGAAATTGCTGCCAGACCTAAAGCTGCCATGGGATGCTCAGTGGCACTTAAAGTGTCTACTGAAATAGTT
34672 TAATAGCTTCTAGAGAGAGAAAATAAAAGGAAATGTTGACTTTTGTAAAGAGACATAGTTGTAGAGATGACCATCCCTGA
34763 AACCATGAAACATTAAGTCTACAGTAATAGAGTGTCTTCAAGGCAACTCACGAATCTGATTTCAAGGGTTGTATTATCTGTTGAAT
34854 TTACATGGCTGTTTATGAAAAGCTTGTCTTGTATTGCTCTCAACAAATTGTTGATGTTGATGTTGACCTTCAATTGTTGACACAGTACTGACA
34945 TCAAAGTAGCTCCCCAGGCTAGAGAAATAGTCAGTCAAACATGTTCTGGTCAGGCCAGGCTTCTGGTTCCAGAACACAGCAAAGCT
35036 GGGTATTGCCCTCATGAAATCTCCCTTCTATGGCTCAAGAACAAAATAGTCATTCTGATTTCTGATTTCAATGCTTCAATTGAAATT
35127 TTCTCTTAATTGAGAAATGTTGTTATAAAAGCTGATAATTAAATCTCATCAAAGCATAAAACACTGATTTCAATTGAAATT
35218 AAGTGTACAAGAAAGTTATAAAAGTTATTGAGATGTTAGCTCTTCTCAATTGTTGACCATGTAATGACTATCAAAATAAAT
35309 ATTATAGTACTACTTTGAAATTCCAGGAAATTGTAAGGTTAGGACTTCATATGTTCAATTACTAAATTATTTTACTCTTCT
35400 TCCTTCTCCATGACTATATTATTTATTTACTTAAATCAATTGCAATTGCAATTGACTTTATATTGTAAGCCTTACT
35491 GTTCTAATTTCACCTAGAATTGAGTTGCTAATAATGACATGCCAAAGTGAATCATTACACAAATTCACATT
35582 ATTCGGACATGGGGCATACAGCTCTATCTGTCACATATAATTATCATTGATTCCTTCTTGTAGGAAATTATGAAATACTTATTCT
35673 GTCCCTATGAACTGTCACATTCAACAGAGAGAGCTAATATAACTAAATTATCCTGTTAGGTTAGTATAATTGAGAAATACTGITA
35764 TAAAGAGGCCAAGATATTGAGGCTTATGGTCCAGGGCTGTGGCTAGGGAAAGGTGAGAAGGGTAAAGGAGGCCAGAAGAAACT
35855 GAAGTGTAGGGCTTCTGATGTAAGGAGGCCAATGAGTTAGGTGTTGTCAGCTACAGAAGAGAAGGCCAAATTATAATGTTGATGAGT
35946 GAATTCTACTCTCTCAAAATGGGACATACAAATCAATTGGAAATGTAAGCTGCCAGTGAAGGACACCCACAGACTGACAGAAGG

36401 GAATTATGGGTGATTCCTCCCTAAATCAGGATGCCCTACTTATAGAAGGAGGAAGAATTAGACAGCATTAAACAACAGATGTCCAG
36492 TGTGATTTGCAAATGCTCTGCTTAAATTCAACCTTGTCTGCCAATGAATAGAGCTTGGAAAAGATTATAAACTAGAGA
36583 TAAATAAATATGTGCAAGGAAAATAACTTGTGAGGTCACTGAATTCCAGGAAACTGAGATCACTGAATTCTGTGTCAGAGTGC
36674 AATTCCACAACACTGAGATACGGACACATCTTAGATACTGCTGTACTGTACCTCCCTGATCCGTAGGCCAGAAGACTG
36765 TTTCTCTGACTACAGCTAATGAGGCTACATAGCTTAAATCTGTTCTTAATGTGAAAATTGCATACTTAAATAGTATTAGTATCT
36856 TGGTTATTAAATGCTTCACTGAATTCTTGATCTCTGTGTCAGAAACATCATAATACATAGGGCAGGTTGGAAGAAGACTG
36947 GCCACAAAGGCTTGAGAGCCTCTCTTAAACTACGTTACAGTATTGCACTGAGAGATAGGGCTATCTATGACAACATAG
37038 TCCTGACTGATGCTAAGGTTGATTCACTGATCTGTAACCGGCCAGAGCAGACAGCTTGTACAGTTACAAGCAACTCTGATCAGT
37129 TAGTAGCTGACTGGAGAACTATGCTTAAAGAATTGAGACTATGCTCAAGCTGTTGGAAAAGTGTACAGTTGATTAGTTGCT
37220 GCACTGATTACAGCAATAGGAAGGGTGCATGTGCACTGTTGTAATCCCTAAACTGGGAAGGTTCCCATTCCTGTTTCA
37311 TATGCAATTCTCATGCTGTGAGCTAGAAGAAAATGATTCTGACCTGTACATATTCACTGCCAGGGCAGTGTAGGGTGAAGAGG
37402 CACTCACCCCTCAGGGTCAGCGGGTACAAGATCAGTACTTCATGCCCTAAAGCAGTACCTCTCTAAATTGCTTGGGTTCTCAT
37493 TTGGTTACCCCCAACCATGGTCTGCACTGCTAGAGGCTTAAACGCAATAGTTATGTAAGGAAACAAATGCATGGAAACAAA
37584 TGTTCAAGAAGAACAAAACACACACAGTAACTGCTGAATGCCATGAAACTCTCTTAATGAAGACAGCCTCGTTGCTGTA
37675 TGTCACTGCTGTTATGAGTCACCTCAGAGTAGACAACATACTTCAGAAAAACACCACGTAACTGAGGTCACCTCGGTAACAGG
37766 GAGCCTAGTTAATGITAATTGGGTTTGTGCTTTGAAAACAGGACACCAGCCCTATGCTCTTAAAGGTTGTTCACTAAAGTAAC
37857 CTGTTGTGACATTGAGGTAAGTGTCTTTATACAAAATCTCTAATGGTTAAAAGAAAAGCTGAGGTTGAGACAGCTGCTCAGTC
37948 GCTCTCTAAATGAATGGCAGACAGATACTCTGGGTAGAATTACAGACCTAGTTAGTCAGGCTTGGTAAGGATCTGCACACCAGC
38039 TTCTCGTTCCCCATCGGGTTCTGTGCTCTACTAGTCGGTCGCCCTGTTCTCCAGGCTTAACTGTGGCTCTTCAGAATAT
38130 TTCTCCAATGTGCTGGAGAGTCGTGCTCTGAGGACACCCCTGCACTGACGGAGATGGGTGCGCAGGCCAGTCTCACCAG
38221 AGTGGCTGGTAGGCCCTCTGGAGCAGGCCGGAGCTTCAGCACGGTCACTGCCAGAGGGCATTGGGCCCTGGCTCCCATACTC
38312 CAGCTGGACTTGGGTGCTGGAACACCTGGTCTTAATGGCCAGTCACTGCCACTTCCGAGGACACGTGCCAGGGTGTGGGGCAGGG
38403 GATGGGCCGGGGAGGACTTGTGATGCAATTGCACTGAGCTTCCAAGGGAGTGAGATAACCTTCATCACAGTCCCAGTGA
38494 TCACTGTATGCTATTGGTGGGACCGAGTTCACTGATATTATTTAATAGGGTGGAACTAAGCACAGAGGGAAATGGCCT
38585 GCCCAGGGTACACAATAATGATGAGGCACTGTTCACTCCCTGTTTCTCAGAGAGAAAATAGGGAGGAACACTGGGAG
38676 GAGAGAGGAGGAATACACAGACAGTGTCTCCCTCTAGCCACTGTCAGGAGCATCACAGCAGGACAGCTTACAGAAATG
38767 TGGGCACAGAAACACTGAGACTCTCTGGTTAACGTTAATCTGGATCTAAACACTCTCTAGTATATACTAGAAAATATA
38858 AAGTCATTGAGATTAGGGCAAAGAGGAAACACTCTGTCTATTCTTTCTTTGAGACAGAGTCAGCTGTGCCAGGCTGG
38949 AATGCGAGTGTGCAATCATGGCTACTGCACTGCCCTCACCTCCAGGGTCAAGCAACTCTCTGCTCAGGCTCATGAGTAGCTGG
39040 AGGTGTGCAACCACACTGCCATTTTTGTTAGGGAGACAGGATTCAACCTGGCTTGGCCAGGCTGGTCTCAAACCTCTGG
39131 CTCAAGTGATCTGCCATCTGGCTCCAAAGCTGGACTACAGGTGACCCAGGCCAACCTGGTCACTTCTGTT
39222 CTCTCTCTTACATTAAAGCCAAGGAAATTCAGTACTTTACTATATTAGCTGACCAATTGGTTCTATCTATACTCATCC
39313 TTATTTCCAGTTTATTCAGCTTCAACTTCTCTCAATTCTTCTTAAATTGTTTAACTGCTCTCTTATCTGCTTT
39404 CTTTCAAGATTGAGAAATGCTTAAATATCTCATTTCCTCAAAACTCAACAAATGAATTAGAATCTACTA
39495 ACATTAGCATCTGGCTAGAGGGACCTCTGATGAAATTAAATACTAAAACGCTTCTGAAATTGCTTAGTCCCTGCTAC
39586 CTTCTCTCTGTTTCTCTGTTTGTGTTTGTGTTTGGAGCAGGCTCTGCCCTGTCACCCAGGCTGGAGTCAGT
39677 GATGCGCCCTGGCTACTACAGCCTGACCTCTGGCTCAGCTCCACCTCAACGCCAACGCTGAGGTCACAGGAGCATGCTAC
39768 CACACTGGCTGATTTTAAATTGGGAGATGGGTCTCCCTATGGTGTCTAGGATGATGAACTCTGGTCAAGTGC
39859 CTGCTCAGGCCAAAGTGTGGATTACAGGCACTGCCACTGCACTGCCAGTATTTTTCTCCGGAAAGCTCTCCCTAC
39950 TTATGCCCATACAGGACTACTTACGGAGGTGTAGTCAGTTGAGGCTACACTGCTCCAAAGTGTCAAGATA
40041 GTGAAATCTGTGATACAACAAATTACTCTCAGCACATTGTTGACAATGACCTCTGGTCTTCTTAAGTTCCAGTGGATT
40132 AAATTCTCTGATGCTCTCTCTTCAAGGAGGCTTATAGAGACAGTTAATGAGGTCACAGCTGGCATCCCCATCTAGAA
40223 GCGCATCGAGAACTCGGAAGGCTGACACTCACTCACAGCAAGCTGGAGAGGCCCTGACAGCATGTTAACAGG
40314 GCTGGCAGGTCTGTACCTGGTGGCAGGCCACCTGCTTACAGATGCTTAGGCATCTCTCTCAGGGAGTGATTATCTTAGC
40405 ATTGCTTTGCTCTCACCTGTCAACAGAAAAGGGCTGAAATTCTCAACAGAGGCAAAATCCATATGTGAAAAC
40496 TTACTTATAACCAGAAATGTGAGAAATTAAATGTAATTAAAAGAAGTCCAGAAAATCTTCACTGGGATTCTTTGTTATT

49051 TCATCACCAAAAAATAACAAATAAGAGTTATGGATTATGATCTACATAAAAGTTTCCCTTGCATTATAAAGCAAAGTA
41042 GAAAGGTATTAGGTGAGATCATGAGGTGATGTAATATTGATAGTTTCCATAACTCTGTGATGCTTTCACAGTTGGAATTITA
41133 TATGGTGAATTTATTTGAAGTCTGTGCAAAATTCAATCAAGGTATGTGTTCTATCACCTTCCAATATTAGTAGTTAATCT
41224 AGTAGATAGAGAGTAGTAGAGTTTCAACATGAAGTTAGCATCTTGACTTGAGTAATCTAGCAATGACTGAATCACCCCTAGATA
41315 ATGGTGAGGCCATCCTTAGGTATCACTGGATGGCACCTGGGCTTCTGTGATGTGGACTACTCAGCTGAGCCAGGGCTCAACCT
41406 GCCAGTGAGTGACATTTGGTCCATAATACACAGAAAAGTGTAGCATGTGTTAGAAGCAAAGCTGAAACCATGGAGAAAAGAGAACAG
41497 CCTAGAGAAAGGTAGGACAAAAAGAAACACAGGTTATAAGGCAAACACTAAGAACTCAGTCAAAGGTGCTCCCGATGGGATTGAT
41588 GAACAGGAATTGTTACAAAGGCCCTGCAATTACCAACCAGGACTGGGCTTGTGACTGTCACAATAACCAAAGATCCTCCCTA
41679 GGCTTTCATTGCTTGAATTGTGCCCTCTATGAGTGTAAAACCTAAACATTGTTGCTCTAATGCCCTCGTACACTCTCCATCCAG
41770 CCAAAAGACAAGAGGGTTATCCTACTTGTACTCAAAGAAGCCATTGCCAGGGTGGCTGGTGTAGTGGCCCAAATGTGTTAGCACA
41861 TGAGAAAAAGACTAGCAATTGCCCTAGTCCACTGGTCCATATTGTTCCCTACTTGAAGAATTAGAGGTATATAATTATCTTCTT
41952 GTGCACTGTGATTGCCACAAATGGAGGCTGGTACTTCAAGGTCCAGGTACAGAAGGCCAGAAGTAAAGAGTGTGATACCCCAAG
42043 GAGATACCTCACCAAATAATGATGTTAGCTGAATTAGGCCAAGCAACTACATCTTATAGGGAAAGCATATTACAGCTGAGCTTATT
42134 TTATTTTGTCTTAGGAAATGCTGAACCTTGTCTTGTGATTAGCTCAGAGGAACCTCTCAATAATCAGTGAACATCATTCTACCTGCACTTGCTC
42225 CAAACTTATTCTACTTCAAGAAGACAAAGAGTGTGATTAGTAAATAACCTTATAACTGTGTTCTCTTACCTAGGCTTTAT
42316 GGTCAATGTTTGGTGCAGAAATTGTTGGAAGTGTAAAGAAGCTCCTCTGTGGACAAACCAAGTGGATCTAACAGGTATAACAAA
42407 ATTACAAAACAAACATCACCTCAAGCTTAAACCTTCTGTCTACCTTAAAGAACATCTCTGACTTCAAGTCAACAGTCTCCAGTCAAC
42498 TTAATGCACTTAAAGCACAACCTGTGCTAGAGTCAACTTGTGACTCTGGAACTGCCATTAGGCCCTCTACATACATCATTAAATTCTACGCAACGCCCTGG
42589 AGCTTACACTTAAAGCACAACCTGTGACTCTGGAACTGCCATTAGGCCCTCTACATACATCATTAAATTCTACGCAACGCCCTGG
42680 AAGTAGATTGTGATTCTATTAGAGATGAGAAAAGTGAAGACATGGAGAGTCAAGCAAGTTCAAGGTATGCCAACAGCAGAGGCCA
42771 GTACTCAGACTTGGGTTCTGTGCTTCTGACCCCTACTCTTACAGTCTTACTGCCCTTTTATAACCTTAAACTCTCCATT
42862 CAAACTCGACAACACTTAGTGGCTCTCTTGTGCTCAAGTAAAGGGTAACTCTTACAGTCTTAAAGAAAATATAGTCAAAGGAGGAATCATAATTAG
42953 AAAACGGCTGGTTTGGGTTCTCACTTCACAAGAGGGTATGTTCAATTCCCCAAATGGGTGATCAGTCTGATGCTAACACG
43044 TGGAGGTGATCAGCTCCACAGGTAAAGGCTCAGTCTCTCTCAAGACTGCCCTGACTTCAGTGCCTCAAGAGGGACCCCCAGG
43135 CCAGCCACGCCCTGTGATCAGGCCAGTACAATTGGGAGTTCTATAACCTGTGACTTCAAGAGAAAACAAGAAATAATA
43226 ATAATAATAATAATAATTGGGAGTTCTATGATACCTGTGAGTTAAATAATCAGTCACTACAGAAITCCAAAAGTACT
43317 CTGGTACTATCAGAGTATTATGATAAAGGGTGAACATCTTTGTGCCCTTATTATGGCTCTGGCTCATCAGTGTGTTCTCAACCA
43408 GGAAGCTCCGCAAGTCTCAGTGTGCTAGAGATTGTTGGGTTCTATTAGTGTAGCAATTGAATACAGTCTCCAGGCCCTCTCTGTCT
43499 CCAGAGGTCTCCAGTCTACCCCTGAACTACAGAAATTGGCTTTTGTGGCCACTTGCATCTGAAGTATCAGGGGCCACAGATG
43590 AGTCACCTCATGGCATCACAAGACACCCACTGTGAGGAATTCTAAACTGTGTTCTAAAGCTCTGTGCTAGGAAATGGAGACAAAGACA
43681 GACATATTCTTATTATACACAAAGCTGATGCTGCCACTCAGCCAACTCACAATGCTGAATGAGCTTCTAAGTTACTGGAAATAGT
43772 GCAAGTGCAGGTATCATTAGGGCCCTGGACAGAGGCCATTACCAACTTAGCAAACCTATAAAATGAAAGGTCCAACACTCCAGTCC
43863 ATCTCCAGGATATGAGAGAGATAATGGAGAGGGGAAGATTGCCAAATTGAAGAGTTGTTCTACATTTCAGAATGCTTCTC
43954 ACTTAAGACACATTCCCTAGCCTGGCTGAAAGCACTGTGTAAGAGTTAACTAATCTCAGACACACATGCTGGAGATGGAG
44045 TTGGCGTGTGCCCCACAGTGTGATCTGTACATAGCAGGAACTGTGTAAGAGTTAACTAATCTCAGACACACATGCTGGAGATGGAG
44136 ATTCCCAGGGTAGGTTGTGAGAGTTGGGAATGAGACTCTGAAACCCAGAGTTACACCAACTGAACTTCAATTACTAAGTGTGATTCTACAGTCT
44227 GAGGTCTAGTGACCCCACTCATCCCTTCAGTGGGTGAGTGTGCCAGCATCTGAACCTATGGTACTCTTCTTCAAGAATGTCGCTT
44318 TAATGAGCTTATTGCTGTGTTGGGCTGGAGTGTGAGTGTGTTTAATGAGGCCCTAGGAGTAAATTCTGAGTT
44409 TTAGTGTCTATGAGTTGCAAAACACATGACTACCTCTCATACCAAGATATAGAACAGATCCACCAACTCCTTAACTCTCCGTC
44500 ATTGGTAGTCCATGTGCAATAGTCTCTTCTGGCTGAGTAGTATTCTGTGACTACCTCACCATTGCTTGTAAATTGAGACATAAGCT
44591 CACTGGAGTGAACATTGGGTTCTAGTTGTTGAGTAAATAAAATGTTGCTGGAGACATTGTTGAGTACAGGTTTGTGATGGACATAAGCT
44682 TTCACTCTCTCTGTACACACATAGGAGGGGCTGGGCTCAATGGTAGTGTGAGTTACTGCATAAGAGACCCCACTGCTTCTCTG
44773 CAAAGTAGTGTGCAATTGCACTTCCACCCCTGTGATGACAGCTGACTGATCTCATCCTGGCCAGCACTGATATTGTTAGTT
44864 CTTAGTCCGCCATTCACTGCTCCCTACATGTGAGGTAGTCATTGTCATTCTGTAAGGAAAAGTTGTCCTTGTGAAATTAGCAGTTG
44955 CAAACTCTCATACCAAGGCTCCACTGTGGATCAGTTGCTGCTGAGGAGGTGGGCCCTACTCTCCCCACAGGCCACATT
45046 CCCCTCTCCCTCTGTGAGCCCTTGTCTGGGGTGGAGAGGAGGAATTCTGCTGGCCACATTCTCCCTGCACTGTAACCTTCT
45137 CCAGCTGGGAGGGGGGTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCTGAGCCCT

45501 GGATTGGCCACCACTGGATGAGTGGGCTGGCTGCCCTCAGGAGGACCCAAACCTTTACAGTGGATTGGACTGAAAACAAGGAGG
45592 GATGCACTTGAATAATGATTATGGTTGCTCTTCACTCTCATAGTGCAGAAAATCCCATTTAGGCAGTCACATAATTATAGCACTAT
45683 TGTATATAAATTTAAAGTTATCTTAGTATAAGTGTGTTATGTTAATTAATAATTTAATTTATTTCAATTAAAGCTTAAAGCTTAAGGGTGA
45774 TAAACTGTACAACAGATATAACCTTTAAATATTTATATAAAATTTCTATATTCAAAGCTTAAGGGTGA
45865 TGCTGTTAATTGGAGACGGGACCTGCACGGGCGAGCCCAGTGAGGGTGGTGTGCTGGAAAGGGCGTGTAGTGA
45956 TTCACTGACTGCTCTAGCTAGCTCTCTGTGATTAAACAGCTAGCTACCTTCAGTTCTCATCTATCCATTAGCAGGAAAG
46047 ACCTCTAACAGTACAAGATTACTCATTTAATAAAAGGACTGGTCAAGAAGGTCTCAAACACAGAATTGTC
46138 ACACACCATAAACAACTTATTCTTAAACACATATAACATTATTGTCCTTTATGCAGGGCACAGACTCTCTGTGA
46229 CTGTTGGTTGCTGATCACAACTCAGCATTTTATAAAACCCACACCCAGTAGCTTGTCCATGATTTCAGTTGCTTGTGA
46320 CAGAGTGAGAGCTTAAAGATCCTTGTAAACATTGAGAGCAGAAGCTCTGGATGTTATGATGTTTCTCCCGAGACTTGTACAG
46411 CAGTCTGTGCACACCTAATGACAGGAATTTTATAGCAACTCATTCTATAATATCTGTC
46502 AGAAATCTTTCTTCCACACCCATGGTCATCGTTCTCATTATCTAATTAGATTGGTCATTAAATAACAGTATAACAGGATA
46593 ATCAAGTTGGTAACAAACAGCATGATTGTCGAATATACTCATCAGGCAAGACAGAAATAGCTGACTAACGGAGACTCATC
46684 CAGCAGGTAGTGTTCAGTGGTAACACACATCAGGGTGTGGCTTCACAGAACGCTGAAGAAAGCTGAAGCAGTGA
46775 GTGGGTTAGAACCTCTGCTGATCTTCTACATCTTGGTTGGACTTTGAACACTGAGACATTCTAGGCATAACACAGATAAAC
46866 TCATGTTGCCAAAAGCATTTCAAAACCTTCTTCTTCATCCCTGCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCT
46957 TCTGTTCT
47048 CTGAGCCTCGACCTCTGGCTCGACCAATTTCCTACCTCACCCCTCTGCTAGCTGGCAGTACAGCTTGTGCCCCACCTGGCAGCTA
47139 ATTAAAAATACAGAGAGATATAATTGTAAGAGACAGGGCTTACTTGTGCCCCAGTCTGGCTCAAACACTGTTGGCCTCAAACGATC
47230 CTCCCTACCTCAGCCTCCAAAGGTCTGGGATTACAAGITGAGCCACTCAGCACCAGGCTGAAGTTTCTTGGCTTCTGTC
47321 ACCTCAACAAACAGTTCTCAGCTGTTCTGCTAGGGCTGTAAGAACAAAGTTCTACAGACTGGGGCTTAAACAACTGAAATTACTGTC
47412 CAATTCTAGAGGTAGAATTGAAATCAAGGTGCCAGCACTGTTGGCTCTGAGGGCTGAGGAAAGGGCTGTTCCAGGCTCTC
47503 TCCCTGACCTGTCAGCAGCCCTTCTCCCTGTCCTTACATCATCTCCCTCTGTTGCTGTCAGATTCTCTCTCTGTAAGGA
47594 CACCAAGTCATATTGACTAGGGCTTACCTCTGCTGGCCTCATTTAACCTGATTGCTCTGTAAGAACCTATCTCAAAGAAGGT
47685 CAGGAGTACTGAAGGTAGGACTTCGACATATACTTTGGGGAACACATTCAACCTATAAAATTCAAGAAAAGACTCTACCC
47776 GCAGAACTTAGCAAATAGATTGATTGACCTTAAAGAACATTCCATTACTGAAATTCCATTACCCCTAGTTGGAGAAGGCACAGGT
47867 GATATTGTTATGATGGGGGAGAAATCTGAGTGTGCTCTACTACAGCTTCTGCAATTGTAAGTGTGAGAACATGAGGCTGTG
47958 CGGGGCTCTGCTGTAATCCCAGCACTTGGGAGGCCAGGGCAGGGGATTGCCAAGGTTGGAGTTGAGACCGAGCTGGCAACAT
48049 AGAGAAAACCCATCTACTAAAATACAAAATTCTGGGTGTTGCTGCACTGCTGTAATCCCAGCACTTGGGAGGCCAGGG
48140 GAATTGCTGAAACCGGGAGGGAGATTGCACTGAGGCAAGATCATGCCATGCACTCAGCTGAGACACAAGGCAAAACTCATCTC
48231 AAATAAAATAAAATAAAATAAGTAAGTGTGAGTAACCTGCTCAGTAATGAGAACGCACTGAGGCTGAGGAAACATTGAA
48322 GATAGACATATTGATGCAATTATGACCTCCCTCTACTACCTGTCGATTGTCGAGGCTGAGGAAACCTTACAAATCCCC
48413 AAAATGGAGGTATAAGGGTTGAGAGGAATAAGCTAAATGTAAGTCTGGGAACATAAGGAATCAAGAAAATGTTAGGG
48504 CATCTTCTTTAAACCTGTAAGGTATTAAATTACTCTGAAAGTCTTCTACGCTGCTCTCTCAATTCTAGGTGCTCTGTC
48595 TAGCTCCAGGCTTGTCCATGCTTATGAGACAGCACAGCTTACGACTTCTGCTCATGTTACGTTCTCATGTTGCTAGTTG
48686 TTGCCATTGAGAACATGAGAAGAAGACCTCTGAGTTACGAGTGGCTAATTACCCAGGAGGAGCTGATGCAAAGATGCT
48777 GCAAGGCTCTGGGAGCTACTGTAATACTGAGAACAGGAGGCTGAGGCTGCTGAGTGCCTCTGGTTCTTATATTAGGTTG
48868 TCTGACCCCTTCTCTGGGTTGATGAGGACTTGTGATGAGGACTTGTGAGGACTTGTGAGGACTTGTGAGGACTTGTG
48959 TTAGAAGACTTTAGGAAACCTCTCCCTTCACTGTAACAAACCCAGGTAAGAACATTCTAGAGATGAGTGGAC
49050 TACCACTTCTCAGAGGACTCTCCACTCTGGGCCAGACCTGACAATATGATGCAATCTGGAGTCATGAGAAGATGTTAGT
49141 CAATTCAACGGAGTAATTGAGGCTTGTGAGGACTCACCTACAAATGCCATTAGGAGAGAAGGAAAGGACTGAGAAC
49232 TCACTGAGGAGTGAACCCAGCATGCCAACGGACCTGACACCCAGGAATGACCTGGACCTGACTCTTAACT
49323 ATCTTCTATCAGACTCTGAGAACGGTCACTCCAAAGTCTGAGCATGGTCTCAAGCAGTCTCAGTAAGGAGG
49414 TCTGTCAGTGGCTCTCTTAATTCAAATCTAGATCTGTTCTGTCACACTCCACGGCAGTTACGTTGAGC
49505 CATTCAACGGTAAATTGGATTCCCAAAAGTCTGAGGACTTGTGAGGACTTGTGAGC
49596 CCAAGATCAGGAAAAAAATGAGGAAACGTTGCTTGTGAGTGC
49600 CAAACAGGAGCAGAAAAGAGAGAAATGCTTATTACATC

B

59151 GAAGGTGTTAGCTGCTGCTACTTCTGAAACAAAGACTAAGGAGAGTTAGCTGATTTACCAAAGTGTCCGAGTAGACTTCCCCA
59242 AATTAGCTGTTAGCTTGGCAGGACTCATCCCACCTCTGAGCTGCCATGTTATAAACTTGGGCTTGAAAGTCATCCTGAAACTGATC
59333 TTAACTTACACATGGGTAATGTAAGTGTAGTTAAAGAAGCAATAAAAAAATAAAGGGGTGATAGATAGATTCAGGATTCAGGAACTGGC
59424 CATATGCCCTCCCTCTATTTCTCCAAAACCCTTAATTCTACACCTAGCTATCCCTATAATTGCTCAAITAATGACAAACATAAAA
59515 GATAGGTGAAAATCCAGTTAAGGGAGTTCTTCAGAATCTGTGAGGTTAAATCAAGTCCTCTCTATAGCAGGCTAATTCCATGATGAA
59606 GCTGTGGTGTAAAGTTTATTTGCTTAGGCAGGTAGCTTCTCTGACATGCTATCTCTCTCTTGAAGTGTCACTCTGAAATGTT
59697 GTTCAGTGTAAAGGGAGTTGATAACAATGTCGGTAACTTAGTCCTGAACTCTGGTACACTGTAGGAAAATAATGTAAGCC
59788 ATATTGATCACCCACCGTCTTATTCAATTGAGGATAATAGTTATTACTAATTATGCGTGTAGGAAACATACAGCCATTATTCTGTCC
59879 CTTAAAGTTTATTGATTGTAAGTCCCTGGTAGAAGAGTACAAAACAAAAGGAAACACTCATGTTAAATAACGTGATTTCT
59970 TCTTGTGTGTACCGTACCGTGCAGGAGACAGGAAAACACCGGTTAACGGTCCACAAGCCAGATAACCATGATGAGTCTAAAGATT
60061 TTCTCTGCTTTAATGTMCTTCAGTTGTCATGAGAAGACTGGATATAGCTCTGTGTTGTCATTAGAGGATTTATATCTCTTCAC
60152 TCTAGAAAATTACCAAGCTGTTAICTCTTAAATATTGTCCTCTCATTTCCATTCCCTGGAGATTCTATTAGATGTGGGTTGGATCT
60243 TCCCTTAAATCTATGTTCTTAACCAATTTCATATTTTTAAATCTCATGTTGATTTCTGATCTGTCATACATT
60334 ACTTCTCCAATTATGTCAGTCTATTATTAACCTTACCTTCAATTGAGTTTTGTCCTTGTGCAACTGCAATTTTTTTTTT
60425 TTTTTTTTTGGTGGAGACAAGGCTGGTCTGTCACCCAGGCTGGAGTGCAGTGGCATGATCTGGCTACTGCAACCTCCACCTCC
60516 AGGCTCAAGCCATCTCCCACCTCAGCTCCAAATAGCTGGACTGCAGTCACCTGCCACCATACCCAGCTAATTTTATATTGTTGTA
60607 GAGATGGGTTTGGCATGTTGCCAGTCAGTCTTAACTTGTGAGGCTCAGGCAATCCACCCACCTTAGCTTCCAAAGTGTGGGATTA
60698 CAGGCATAGGGCACCAACCCGCCCTAGTGCAGTCATTTCCTTTCTAGAAGTCTGTTGGCTCTTCTGAACTTTCTGGTCTC
60789 TTTTTCTATTCCCTGTTATTGTTTATCATGGATTATACCTTTATTATTTCTTTCTAGTAACTTCTGTTTCTAGTACATGATTACTTA
60880 GGATTTATACATTATTACCTCTCATATAATTAACTTGGTTTACATTGTTTATTATCTTAATTCTTATTGGACTAAITCTTTGTT
60971 TTGATGCATCTGTTAGTTCCCTCGGGGGTGGTTCTCATAGGTTGTAATTTTTATTGTGAGGCTCATCTGGGAGAGTGGC
61062 TCATATGCCCTGATTGAGAATGTTCCCTCCAGAACAACTTTATGTTGGTTGGCTGAATCCTAGTAATCTGGACTGGCTT
61153 TTAAGTTATTCTCACCTGAAAGCACATACAGTCAGGAATGTAACATTGTAACTTAACTATGTTGCTGTCAGGCTAGAATTCTCATT
61244 TCTCAATATGACTTTCTTCCATAAATGGCCAAGCTGATAGCAAGTATTCTGCTCTGGACATCTGGCAGATTCTTCTA
61335 CCCTCTTCATAGATGGGATAGCTTCAAGGCTCTGGGAGATAGCAGGCTGATACCTGAGAATCTCCCTTTGCCCAAGGCCACA
61426 TCACTAATCTTGTGAGCTTGAAGCCCCCTCCCTCAGGCCATAGACCTATAACAGCTCTAAATCTTAATGGCAGTCTACTGACA
61517 GCTGCCCTGTCACCGCTCTGATCATCTAGCTTGAATTCTTCTTCTGGCACCTGGTGTGTTCCCTTCTTCTAACAAAGT
61608 TTTTCTGCTTCTAGCACTTCCATATGTCAGTGGAGGAGGCTGAATGCCATCTGCTCTGCTGTCATGTTGCTGAAATCACA
61699 GTGAGTTTTGTGAAAGTGAACAGCTCCATTCTGAGCTGTTGAGTCTGACTCTTAGATCCATCCTTCACAGTGTATCTGG
61790 GCAAGCATCTAACATTCTGAGTTGAAAGAAAATAGAGATAATGCTGACTTTTAGGGTTGTTGGAAAATTAAACAGATAATGCA
61881 TGTAACAACTTCTGAACTTCTGCCACATAGCAACTCAGGGTCTCAATTCTGGCTAGGTTGGCTCAGTATCCCTGGGCAACTT
61972 TTAACATAGACATTCTGGGCCAGAAGCAATGCCACACTTGTAGTCCCAGCTACCCAGGGCTGAGGTAGGAGGACTACTGGGGCC
62063 AGGAGGTCGAAGTTGCCATGAGCTGATGCCACTGCACTCCAGGCTGGTGAAGAGTGGCCCGTCTCAAAAATAATAAAATA
62154 AGGTAAAAAAACAAAACACAGAGATTCTGGGTCTACCCCTCAGCAGTTGATTCATTAAGTCTGAGATAAGATCCAGAAATCTG
62245 CATTGAAAAGCTCACAGGTGATCCCATAATGCCACCCAGTTGAAACAGTCTTAAAGTTATGCAAAAATGTAATTCTGAA
62336 TTCTCAGATTCTAACACTTCAAGTCATTATTCTCCACACTGTGATATTCTCATCTCAGATGTTGAGGAGGAAACAGC
62427 GTCTCATACGGCAGACCAGAGGGAAATCAGCAGGAACCTCAAAGAAGACTATAACAAACCTCAAAGAGAACCTCAGGCCATGAGCG
62518 GAAAATTCCAGAACTGACAACCAATTCTAGAGTTGAGAGTCAAAAGAGGTAAGAACAGGGCAGAGGGAGGCTCTCCCTGTTGGATAAA
62609 GAGCACGGCATGGGCTAGCACCTGGGCTAGCTCTGCTGACTTGGGAGCTGCAAGAACCTCGAAAGGGTGAAGAGGGTCCCAAGT
62700 CAGAGGGCTACAGAGTGTATTCAATTCTGCCCTGCTCTCCCATCCCTGCTCTTGCACCTCTCCAGACACCTGGTGTGCTTGT
62791 GCCAGGGTATTCCACGGGCTGAATGATGCCCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGCTGAGG
62882 GGTGCGAGACCTTCACTTATTCTGAGTTGTCATGACTGATGCCATTTCTACTGGGTGATCCACCCCAACCCCTTCTAAAGG
62973 CTAACTGATCTTCTGCTCTGACGCTCTTCCCTCTCTCTTCTTCTTAATTCACTGGGACTCTTCCACAGATCTAGTT
63064 CAGGAATGTGAAACCCAGTTGTCACAGGGCAGCTAAGAAAAGCCATCTCATTGTTGGAGACTGTGGCCCTGCAACCCCTGGAGAAGGACT
63155 TGCTGGTACTAAAAAATGGGACATTGCAACCCAGGACTGACTGTACACTCCCTGATCAGCCAGCACTCTGGAGCTTGGGATCCAGG
63246 AACCATGGAATTATTCCAAAATGGACTCTGACCAGATTGCTACTGGGGGGTGGCAGGAGTGGGACTCTGAGGATGGTACTCAGGATGACTGC

63701 TGTTTTATTACTGTACATTAATTTAACATGCAATTATAGAAGAACATCAGTAAAGCACTGATGTAGGAGATAACGGTACTTGGACC
63792 AGTCAGGCCAGAAAATCAGAATCTGCTTCACTTAAATGGAAACAATTCTCGATAATGCTTGCCTTTCTTATGCACTCTGTGTA
63883 CTATCTATTTCTCCCTCTGGGACCAAGTTCTTCTTATAAAGCAATAATATCTGTTTCATTCAGAACATTGCTGCTGTCAG
63974 CATATGATATCAGCTACAAAATATTCACATTGACTCTTGAACAGGACTTTAGGAAAAGAGGAACAAAGACATTATGAGAA
64065 TAAATATATATTTAATATGACTGTGACCTTGACTGATAATAAAGATGATAAAGAAATTGCAAGCTAAATGTTTCCCTTGCAACTCA
64156 TGCTTTGTGTTTGATGACCTACTCGCTGTATGTTGTAAGGCACTTCAGAGAGAAGACAGATGCATCATCCTGGCCCTCCATC
64247 AAATAACACTATCCAAGGTGGCACCTCTCGCAATGTTAACCTGCTAGTAATGACAGTGAATTGCGGATATTTCAGAACATTTT
64338 GTTATACCATCAGGTATGCAATGAAAGAGGACTTAAATAATAATTAAACTTACAGCTTAAGTCTAAACTTTTA
64429 TTTTAGGTATTGGGAAGAACTTTTAAAGTATAACCTAATGCTTTAAATGAGTACACATGACATACTTAAATCCATATG
64520 TATTCCTACTCTGGGAGACACTGTGTTGAGACCAAGGTCAAAAACGTGGTACCGCCCTCAAATGTCCTGGTCCCTGAGGAAG
64611 ATCATATACCTGTGTAAGGCCACAGTACAAAACAGACTAGAACACAGCCCATAGCATGTAACCTTCTGACTAACTCAAGGATAGGCC
64702 AACACCTATGGTATTAGATTGCCCCTAAACAATAAGAGTTAGATGCTAAGTTATAGTCCTGGCACCTAACAAATAGGCCAAATAG
64793 CCCTAGTAACCTAGAATATTCTGATTAATATCCCCTGCTTTAGATAACCTGTTGCTATTGGGTTGTTTACAGTCTTTGTA
64884 CCACAGTGGATACATTGCTCATGAGTGCAGGAACATGTTACTGTCATTCTTACCCCTAGGCCCTGCAACAAACACACAAAAGATAC
64975 CCAATAAAATTTGTGATTCACTAAATGAATGATGAGTAGGCCCTGCTCTAGAAGTGCACTGCCAATAAGAATGTAATGCAAGCCA
65066 CATATATAATTTAAAATTCTAGTAGCCATATTAAAAATAATAAGGCCAAGTGCAGTGGCTCATACATGTAATACAGCAGTTGGAA
65157 GACCAAGGTGGGAGATCACTTGAGGCCAGGAGTTGAGACCAAGGCTGGCAACATGGCTAACCCCATCTCACAAAAAGATATAAAA
65248 AATTAACCAAGTGTGGCATGTGCTGTAGTCCCAGCTACTGGGAGGCTAAGGTGGGAGGATGCTTGAGCCAGAAGGTTGAGGCTG
65339 CAGTGAGCCATGATCGTCACTGCCTAGCCTGGTGACAGAGTGAGACCCCTGCTCAAAAATAATCAGCATCATAAAAGAACCA
65430 GCAAAATTAACCTTACTAGTATATTAAACCAATAATAATAAAATTATTTCATATGCTTCACTATAAAAATTATTACAGTCTT
65521 TATTCATATTAAAGTCTTAAAATCTGATGTGAGTTGACTTACAGCACGTGCACTTACAGTGGCCACATTAAAGTGCACAGTA
65612 GCCACAGGGGCCACTGGCTACCATTTGGATAGTGCCTACTAGAAGCTTCACTGGATGCCCTGATTGTGGACTCG
65703 AATACAGATAACCAAGAAGTGGACTAGTGTCTGAAGTAAGAATGACAGGGTATGATTGAGACCCCATGAGCTTACCTAGGAGAAC
65794 TTGTGGGTTGAGAATAAGGATTGTCATATTGGCTCTAGCTGTTACACTATTCTGGGCAACTCCAGATCATTCTCACTCCAG
65885 ATAGTTAAGTGGGAGCATGGCTGCACTTTAAAGTGTGGCACAAAAAAAGTATTGAACGTTGGCTCTGATTATATACTTAAATA
65976 TGCACTTAAAGAGGCCCTTTAAAGAATCCCTAAGGTAAGCAATTAGTATCTTGTGAAATTAGAAGAAACTTGTATGCCA
66067 TGATACCCCTTCATTTATTGGAAAACCTCTATGAAAGCTTAAITAGAGAAAATTCTGATTACCTTCATGCCCTCCCTTCAG
66158 CAAGGTCAAGGGTGCAGTTGTCACTATCACATAAGAATCTCATAAAAAAATTAACATGAATAACTGCACAGATCTGATTGGGTTGTCATG
66249 CCACACATTGTTAAATTCCATAATTCTATTCTATAAAAGAGTTCTATGACAATAGATGTTTAAACAAACAAACAAACAAAAT
66340 TTAGAGTTGTCATTGTAATTGGGTTGCAAGTATGCTTCAAGGACAGAGCTTGTGTTGATGAAATTGTAATTTTCTTTTCT
66431 TTTTGATACGGAGTCTCACTCTGTCCTGGCAGGCTGGAGTGCATTGGCACCATCTAGCTCACTGCAACCTCCACCTCGTGGTCAAGCA
66522 ATTCTCCCTGCCCTCACGCTCTGAGTAGCTGGGATTACAGGCGTCCACCCACGCCCTGACTAATTGTTGATTAGAGATGGGTT
66613 TTACCATGTTGGCCAAGCTGGCTCAAAATCCCTGACCTCAGGTGATCCACCTGCCCTGCCCTCCCAAGTGTG

B

Putative promoter sequence of human CLASP-5

GGAAACAATTCCTCTCATGTGTATGGCTCCCTAAAGTGTGGCTGAGCATTGTCCACATGGGTG
ATGCAAAGGATCACTGAACTAGGAGCAGTTGGAAAAAAATACAATCATTGGAAATTCTGTAGC
ATCGAATGTGCCTACAGGGAGGTAGAAGTATTACATAACAGTTCTCTGGTGTCTGTGTTGA
GCAACCAGTCAGCCAAAAGGGTTCAGCTGCTGAAATGAGAATGGCTGGATCAAATGGCAGCT
CATGATTAAAGGATTCTAGTCAGATACAGACATCCTCACATAGAGAAAATCTGAATGGCTG
GGGGAGAAGGAGTCAAATGCCCTGGATCTTTCTGGGCCTCAAAGTCCTCCTCTGTATCA
TCCTTCAGTATTGGCAGGACCTGACTGCAGGCATCATGGCCTCTGTGAACCTCTCAAGGGTA
TGTATTATCTGACAAAATACGATGTCCACTAACAGGCCACTGAAAGGTATCTAGTCAGTTC
TGCTCATGGCCCAGCCAAGGCCTACGTTTATAACATGATATCAAAGATTGCATCTAAAATGT
GATGATTCTCTAAAATAATCATTCTAGATTCTATTTAATCCAAGGTATTCTCAGC
GGAAATAAGGAAACAGTTACTCTCCCACCAAAACCTTGGCCAGTACCATCGACAGAGCATAAGT
ACCTCTGGCTTCCCCCTCTCAACTAGTAAGTATGAGTCCAGGTTACTTAGCGATTGGTCA
AGTGAAAAGTGCCAGGGTATGTGTTGCCTCCTGTTAGATCTTACCATCACCTCA
CATTCTCAGTCACCAGATCCTAACTCTGTGACTGTGTGGACATCAGACAATATCCCTCT
CTCTCTGCCAACCGGTACTTAGGGTACATAATAGAACCTCTGGGAGCTGTGGTTTGATGTCTC
TAGACTAGGTGGGCTTCCAGGTGACTCAGTCTCATCCAAATTATGGTTCATATTGGGGAGAA
GGGCTAGCCCCAAAATTAACCAACATTGTAGTATGCATTTTGGAAAAGCATATTCCAAA
TCTGAATGCCAAGTTACAGACCTCTTTGTAAGGAAATTTCTTGTAGTATAATTACAT
ATAATAAAATTACACATTAGGTGACAATTGGTGAACCTGGCAACTTAGAGTCACCTAA
CCTTCTCAGTCAGATATAGAACACTTCTTTATCCTAAAGCGTCCCCAGCGCGCTTAC
AATCTCTCTCCCCAGGCCACACCCCTCAACTCAGCAACTCTGACTCACCTGTACC
ATTGTCTATCTGGAGCTTCATATCCTGTACAGTATGTACAAACCTCTTTGAGACA
GGGTGTCAGTCTGTACCCAGCCTGGAGTACAGAGGTGTGATCTCAGCTCACTGCAACCTCAAC
CTCCCCAGGATCAGATGATTCTCTCCACCTCATCCTCCAAAGTAGCCGGACTACAGGCGCAT
GCCACACACCTGGCTAATTGGTACTTTGTAGAGACAGGGGTCTCGTATGTTGCCAGG
CTGGCTTGAACCTCTGGGCTCAAGCGATCCTCTGGCTCAGCCTCCAAAGTGCTGGATTAC
AGTGAGCCACTGCACCTGGGCTAAACCTTCATTAAAACACATTCTCTTAAATTGAAGA
TTGCCTACATTATCAATGCCAATTGGTGAAGTGTGCTTATATGTGTTATATTGGAGC
ACTAAATGCCAGATGTGTCAGATAATCTGACAAATGAGATGGTTGTAAACCG
AGTGAATATTCACTCTCTGTGAGAGAGCTCCAGCCCTCTGTACTCACCTCCACACAGCA
CAGCAGCACTCTGCTGGTTCTGCTGCTTATCTGAAGAGGTTAGGTTACTTTGTTCTACT
TATTACTTCGAAACCACCTCTGCCTAGAAATTGTAAACCTCCGCTCAGTTCCGTAACCG
CCATTGTCCTGTAACAATTACCGCCGTGTAACGTGAATCTT

C

FIG 7

hCLASP4	-----MFPMEDISISVIGRQRRTVQ-----	20
hCLASP5	-----MTHLNSLDVQLAQELG-----	16
hCLASP3	-----MAERRAFAQKISRTVAAEVRKQIISGQYSQSPQLKNLNIVG	41
hCLASP2	-----MLLPFYDDFQTAIRRQGRYICS-----	23
hCLASP7	-----MAASERRAFAHKINRTVAAEVRKQVSERSGSQPHSSRRCSSSL	43
hCLASP1	MSFRGKVFKREPSEFWKKRRTVRRVIQEFFFHRFSSQEKPRLLPDLYETVIEELEKTYRN	60
	...	
hCLASP4	-----STVPEDAEKRAQSLFVKECIKYSTDWHVVNYK-----	53
hCLASP5	-----DFT-----	19
hCLASP3	N-----ISHHTTVPLTEAVDPVDLEDYLITHPLAVDSGPLRDLIEFP	83
hCLASP2	-----TVPAKAEEEAQSLFVTCECIKYNSDWHLVNYK-----	55
hCLASP7	G-----VPLTEVVEPLDFEDVLLSRPPDAEPGPLRDLVEFP	79
hCLASP1	DPLQDLLLFFPSDDFSAAATVSDIRTLYSTVPEDAEHKAENLLVKEACKFYSSQWHVVNYK	120
	::	
hCLASP4	YEDFGDFRMLPCKSLRPEKIPNHVFEIDEDCEKDED-----SSSLCSQKGVIQKG 105	
hCLASP5	DDDLDVVFPTPKECRTLQP-SLPEEGVELDPHVR-----DCVQTYIREWLI 63	
hCLASP3	PDDIEVVYSPRDCRTLVS-AVPEE-SEMDPHVR-----DCIRSYTEDWAI 126	
hCLASP2	YFDYSGEFROLPNKVVKLKLPVHVYEVDEEVDKDED-----AASLGSQKGGITKHG 107	
hCLASP7	ADDLELLLQPRECRTTEP-GIPKD-EKLDQVR-----AAVEMYIEDWVI 122	
hCLASP1	YEQYSGDIRQLPRAEYKPEKLPHSFEDIHDADKDEDTSHSSKGAGGTGVFKSG	180
	::* . :* .	
hCLASP4	WLHKANVNSTIT--VTMKVFKRRYFYLTQLPDGSYILNSYKDEKNSKESK-GCIYLDACI 162	
hCLASP5	VNRKNQGSPEIC--GFKKTGSRKDFHKT-LPKQTFESETLECSEPAQA--GPRHLNVLC 118	
hCLASP3	VIRKYHKLGTGF--NPNTLDKQKERQKG-LPKQVFESDEAPDGNSYQDDQDDLKRRSMSI 183	
hCLASP2	WLYKGNMNSAIS--VTMRSFKRRFFHLIQLGDGSYLNLFYKDEKISKEPK-GSIFLDSCM 164	
hCLASP7	VHRRYQYLSAAY--SPVTTDTQRERQKG-LPRLQVFEQDASGDERSGPEDSNDSRRGSGSP 179	
hCLASP1	WLYKGNFNSTVNNNTVTVRSFKKRYFQLTQLPDNSYIMNFYKDEKISKEPK-GCIFLDSCT 239	
	::* . :* . : . .	
hCLASP4	DVVQCPKMRRAFELKMLDKYSHYLAEETEQEMEEWILTLKKIIQINTDSLVQEKKETVE 222	
hCLASP5	DVSGKGPVTACDFDLRSLQPDKRLENLLQQVSAEDFEKQNEEARTRN-----RQAE 169	
hCLASP3	DDTPRGSWACSI FDLKNSLPDALLPNDLDRTPNEEIDRQNDDQRKSN-----RHKE 234	
hCLASP2	GVVQNNKVRRAFELKMQDKSSYLLAADSDEVEMEEWITILNKLQLN-----FEAAMQEK 219	
hCLASP7	EDTPRSSGASSI FDLRNLAAADSLPSLLERAAPEDVDRRNETLRRQH-----RPPA 230	
hCLASP1	GVVQNNRLRKYAFELKMNLDTYFVLAATESDMDEWIHTLNRLQISPEGPLQGRRSTEL 299	
	:: . : . . .	
hCLASP4	TAQDETSS---QGKAENIMASLERSMHPPELMKYGRETEQLNKLRSRGDRQNLFSFDSE 278	
hCLASP5	LFALYPSVD---EEDAVEIRPVPECVKPEHLG-----N-----RILVKLLTLKFEIE 212	
hCLASP3	LFALHPSPD---EEEPIERLSPVDIPKPEHFG-----QRLLVKCLSLKFEIE 277	
hCLASP2	RNGDSHEDD---EQSKLEGSGSGLDSYLPPELAKSAREAEIK---LKSESRVKLFYLDPD 272	
hCLASP7	LLTLYPAPD---EDEAVERCSRPEPPREHFG-----QRLLVKCLSLKFEIE 273	
hCLASP1	TDLGLDSDLNSVTCECTPEETDSSENNLHADFAYLTETEDTVKTRNMERLNLFSLDPD 359	
	.* . : . : .	
hCLASP4	VQRDFS---GIEPDIKP-FEEKCNKRFLVNCHDLTFNILGOIGDNAKGPPTNVEPFFI 333	
hCLASP5	IEPLFAS---IALYDVKERKKISENFHCDLNSDQFKGFLRAHTPSVAASSQARSAFV 268	
hCLASP3	IEPIFAS---IALYDVKEKKKISENFYFDLNSEQMKGLLRPHVPPAATTLARSAIFSI 333	
hCLASP2	AQKLDFS---SAEPEVKS-FEEKFGKRIIVKCNDSLFSNLQCCVAENEEGPTTNVEPFFV 327	
hCLASP7	IEPTEGT---IALYDVREKKKISENFYFDINSNSMKGLIYAHGTHPAISTLARSAIFSV 329	

Cadherin EC motif

hCLASP4	PMAFAKPKLQR-----VQDS--NLEYSLSDEYCKHHFLVGILLRETSI	1060
hCLASP5	FFMNADTAPTSP--CPSISSQNSSCSSFQDQKIAASMFDLTSEYRQQHFLTGILLFTELAA	1085
hCLASP3	PCSLLTTPPASPSPSVSSATSQSSGFSTNVQDQKIANMFELSVFPFRQQHYLAGIVLTLEAV	1196
hCLASP2	PMPPFGKGRIQR-----YQDL---QLDYSLTDEFCRNHFLVGILLREVGT	1052
hCLASP7	PCCPLSPPASPSPSVSSTTSOSSTFSSQADPKVTSMFELSGFPFRQQHFLAGILLTELAL	1119
hCLASP1	PIRSANIPDPLTP-----SES----TQELHASDMPEYSVTNEFCRKHFLLIGILLREVGF	1157

hCLASP4	ALQDN---YEIRYTAISVIKNLLIKHAFDTRYQHKNQQAKIAQLYLPFGVLLLENIQRL	1116
hCLASP5	ALDAESEGEGISKVQRKAVSAIHSLLSSHDLDPCKPVEVKVIAALYLPLVGIILDALP--	1143
hCLASP3	I LD PDAEGLFGLHKKVINMVHNLLSHDS D PRYS DPQ I KAR VAM LYPL PLIGI IMET VP--	1254
hCLASP2	ALQEFR---EVRLIAISVLKNLLIKHS FDDRYASRSHQARIATLYLPLFGLLIENVDR I	1108
hCLASP7	ALEPEAEGAFILHKKAISAVHSLLCGHD TDPR YAEATVKARVAELYPLLSIARDTIP--	1177
hCLASP1	ALQEDQ---DVRHLALAVLKNLMAKHS FDD RYREPRKQAOQIASLYMPLYGMLLDNMPRI	1213
	* * * * *	

hCLASP4	AGRDTLYSCA-----	-AMPN-S---ASRDEFPCGFTSPANRGSLSTDKDFTAYGS	1160
hCLASP5	-----QL-----	-CDFTVADTRRYRTSGSD-----	1162
hCLASP3	-----QLY-----	-DFTETHNQRGRPICIATDD-----	1276
hCLASP2	NVRDVSPFPVNAGMTVKDESALPA-----	-VNPLVTPQKGSTLDNSLHKDLGAISSPYTT	1167
hCLASP7	-----RLH-----	-DFAEGPGQRSRLASMLDSDE	1201
hCLASP1	YLKDLYPFTVNTSNQGSRDDLSTNGGFQSQTAIKHANSVDTFSKDVLNIAAFSSIAIS		1273

hCLASP4	FQ-NGHGIKREDSRGSLIPEGATGFPDQGNTGEN-----TRQSSTRSSVSQYNRLDQYE	1213
hCLASP5	-----EEQEGAGAINQNVALIAAGNNFNLKT-----SGIVLSSLPLYKQYNMLNADT	1208
hCLASP3	-----YESESGSMISQTVALMAIAGTSVPQLTR----PGSFLLTSTSGRQHTTFSAES	1324
hCLASP2	STPNINSVRNADSRGSLISTDSGNSLPERNSEKSNSLDKHQQSSLTGNSVVRCDKLDDQE	1227
hCLASP7	-----GEDEGDIAGTINPSVAMAIAGGLPGRSR-----ASISQGPPTASRAGCALSAES	1249
hCLASP1	-----TVNHADSRSLASLDSNPSTNEKSSEKTDNCCEKIPRPLALIGSTLRFDRLDQAE	1327

hCLASP4	IRSLLMCYLYIVKMISETDTLLTYWNKVPQELINILILLEVCLFHFRYMGKRNIARVHDA	1273
hCLASP5	TRNLMICFLWIMKNADQSLIRKWIADLPSTQLNMRILDLLFICVLCFEYKGKQSSDKVSTQ	1268
hCLASP3	RSLLLICLWLWVTKNADETVLQKWFDTLSVLQNRLLDLYLCVSCFEYKGKKVFERMNSL	1384
hCLASP2	IKSLLMCFLYILKSMSSDALFTYWNKASTSELMDFFTISEVCLHQFOQYMGKRYIARNQEG	1287
hCLASP7	SRTLLACVLUWLKNTEPALLQRWATDTLPLQLGRLLLDLYLICLAafeYKGKKAFERINSL	1309
hCLASP1	TRSLLMCFLHIMKTISYETLIAYWQRAPSPEVSDFFSILDVCLQNFRYMGKRNIIRKIAA	1387

hCLASP4	WLSKHF G IDR-----	-KSQTMP ALRN RSGV M QARL QHL SS LESS-----	1311
hCLASP5	VLQK S RDV KAR-----	-LEE A I RL RGE GARM E MRR RAPG ND RFP G L NEN-----	1311
hCLASP3	T FFK K S K DMR AK-----	-LEE A I LGS I GAR QEM VRR S RQ G L E R P S P G S A F Q S Q-----	1430
hCLASP2	LGP I V H D R K S-----	-QTL P V S R N RT Q P M H A R L Q L G S L D N S-----	1323
hCLASP7	TFK K S L D M K AR-----	-LEE A I LGT I GAR QEM VRR S R E R P S F G N P E N-----	1350
hCLASP1	A FK F V Q S T Q N N G T L K G S N P S C Q T S G L L A Q W M H S T S R H E G H K Q H R S Q T L P I I R G K N-----	-	1442

hCLASP4	-----FTLNHSSTTTEADI FHQALLEGNTATEVSLTVLDTISFFTQCFKTLQ	1359
hCLASP5	-- LWRWKKEOTHWRQANEKLDKTKAELDQEALISGNLATEAHLIILDMQENITQOASS-ALD	1368
hCLASP3	ENLRWRKDMTHWRQNTEKLDKSRASIEHEALIDGNLATEANLIIDLTLIEVVQTVS-VTE	1489
hCLASP2	-----LTFNHSYGHSDADVLHQSLSLEANIATEVCLTALDTLSFLTIAFKNQLL	1371
hCLASP7	-- VRWRKSVTHWKQTSRDVKTDIEMEHEALVEGNLATEASLVVLDTLEIINVQTMV-LSE	1407
hCLASP1	-- A1.SNPKLLQMLDNTMTSNSNE DIVHHHVDTAEIATEGCLTILDLVSLFQTQHQRLQ	1500

ITAM		ITAM			
hCLASP4	EKGTEENVKIIQDSDKVNAKEELDPHYAHIQVTVK	YFDDKELTERKTEFERNHNISRFV	1799		
hCLASP5	QCFGAEFVEVIKDSTPVDKTKLDPNKAYIQVTV	YFDFEYEMKDRVTYFEKFNFLNRRFM	1810		
hCLASP3	ERFGEDVVEVIKCDNSPVDKCKLDPNKAYIQVTV	YFDTYEYMKDRITYFDKNYNLNRFRM	1932		
hCLASP2	DKGSENVKMIQDSGKVNPKDLDShAYIQVTV	EFFDEKELQERKTESHNRIRRFM	1770		
hCLASP7	ERFGDDVVEIICKDSPVDKSKLDShAYIQVTV	YFDTYEYLDRVTYFDRNYGLRTFL	1851		
hCLASP1	DKFGADNVKIIQDSNKVNPKDLDPhYAYIQVTV	EFFEEKEIEDRKTDFEMHHNINRFV	1972		
	: * * : * : * : * : * : .	: * : * : * : * : * : * : .			
ITAM		DOCK motif			
hCLASP4	FEAPYTLSGKKQGCIEEQCKRRTI	LTTNSNSFFVKHRIPINCEQQINLKPIDGATDEIKD	1859		
hCLASP5	YTPPTLEGRPRGELHEQYRRNTVLT	MHAFFHYIKTRISVIQKEEVLTPIEVAIEDMKK	1870		
hCLASP3	YCTPFTLDGRAHGEHQFKRKTI	LTTSHAFFHYIKTRVNTHKEEILTPIEVAIEDMQK	1992		
hCLASP2	FEMPFQTGKQRQGGVEEQCKRRTI	LTAFHFFVKHRIPVMYQHHTDLNPIEVAIDEMSK	1830		
hCLASP7	FCTPFTPDGRAHGEPEQHKRKTLL	STDAFFHYIKTRIRVCHEETVLTPVEVAIEDMQK	1911		
hCLASP1	FETPFTLSGKXHGGVAEQCKRRTI	LTTSHLFYVKHRIQVISQSSTELNPIEVAIDEMSR	2032		
	: * : * : * : * : * : * : * : * : * : .	* : * : * : * : * : * : .			
Coiled-coil					
hCLASP4	KTAELQKLCSSTDVMDIQLQKLQG	WWSVQVNAGPLAYARAFLNDQSASKYPPKKVSELK	1919		
hCLASP5	KTQLQAVAINQEPPDAKMLQMVQ	SGATVNQGLEVAQVFLAEIPADPKLYRHNNKLR	1930		
hCLASP3	KTQELAFATHQDPADPKMLQMVQ	SGTVVNQGLEVAQVFLSEIPSDPKLFRHHNKLR	2052		
hCLASP2	KVAELRQLCSSAEVDMDIKLQLQK	SVSVQVNAGPLAYARAFLDDNTKRYPDNKVKLL	1890		
hCLASP7	KTRELAFATEQDPPDAKMLQMVQ	SGVPTVNQGLEVAQVFLAEIPEDPKLFRHHNKLR	1971		
hCLASP1	KVSELNQLCTMEEVMDISLQLKLQG	SVSVKVNAGPMAYARAFLEETNAKKYPDNQVKLLK	2092		
	* : * : * : * : * : * : * : * : * : * : .	. . : . : .			
Coiled-coil					
hCLASP4	DMFRKF1QACS	ALELNERLIKEDQVEYHEGLKSNSFRDMVKELSDI	IHEQILQEDTMHSP	1979	
hCLASP5	LCFKEFIMRCG	RAVEKNNKRLITADQREYQQELKKNNYNKLKENLRPM	IERKIPELYKPIFR	1990	
hCLASP3	LCFKDFTKRCEDALRKNKSLIGPVQKEYQRELGKLSSP	-----	-----	2090	
hCLASP2	EVFRQFVEACGQALAVNERLIKEDOLEYQEEMKANYREMAKELSEI	IMHEQICPLEEKTS-	1949		
hCLASP7	LCFKDFCKCDEDALRKNNKALGPQKEYHRELERNYCRLREALQP	LTLTQRLPQLMAPTP-	2030		
hCLASP1	EIFRQFADACGQALDVNERLIKEDOLEYQEELRSHYKDMSELSTM	VNEQITGRDDLSKR	2152		
	* : * : * : * : * : * : * : * : * : * : .				
PDZ ligand					
hCLASP4	WMSNTLHVFC	AIISGTSSDRGYGS	PHYAEV	-- 2008	
hCLASP5	VESQKRDS	FHRSSFRKCET	QLSQGS	----- 2015	
hCLASP3	-----				
hCLASP2	VLPNSLHIFNA	ISGTP	TSTMVHG	MTSSSSVV	1980
hCLASP7	--PGLRNSLN	RASFRKADL	-----		2047
hCLASP1	GVDQTCTR	VISKATPALPTVSI	SSSAEV	--	2180